1994

Statistical inference for real-valued Markov chains and some applications

Gregorio Saravia Atuncar
Iowa State University

Order Number 9518355

Statistical inference for real-valued Markov chains and some applications

Atuncar, Gregorio Saravia, Ph.D.

Iowa State University, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Statistical inference for real-valued Markov chains

and some applications

by

Gregorio Saravia Atuncar

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Statistics
Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1994

To the memory of my parents.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to acknowledge my advisor, Dr. Krishna B. Athreya for his expert guidance. From the formulation of the problem to the reading of my drafts I have felt comfortable in discussing the material with him and learned more in each session.

My thanks also to the members of my committee: Dr. Y. Amemiya, Dr. D. Mcbeth,Dr. W. Meeker, Dr. S. Sukhatme, and Dr. A. Weerasinghe.

I also would like to thank Dr. Dean Isaacson for his help and encouragement.

My thanks to the "Universidade Federal de Minas Gerais" and the "Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior" (CAPES) in Brazil for providing me the oportunity to follow my graduate studies at Iowa State University.

I would like to acknolewdge Jun-ichiro Fukuchi for his comments and help with computer work.

Finally, my special thanks to my wife, Ana, and my son, Goyito, for putting up with my long working hours.

1. INTRODUCTION AND PRELIMINARIES

1.1 Overview

Let $\{X_n$: n=0,1,2,...$\}$ be a real-valued Markov chain. The purpose of this dissertation is to study properties of kernel type estimators of the stationary density and of the transition density of such a chain.

An extension will be given to consider semi-Markov processes when the imbedded Markov chain has R as its state space and satisfies appropriate conditions.

The topic of Statistical Inference for Markov Chains started receiving attention in the fifties. Bartlett (1951), considered the problem for a Markov chain with a finite state space. Derman (1956), gave asymptotic results for estimators of the stationary probability distribution and the transition probabilities for a Markov chain with a countable state space.

The case of general state space was considered first by Roussas (1969a) and by Rosenblatt (1970). Both authors extended to the Markov chain case some results for kernel estimators of the density based on a sample of independent and identically distributed (i.i.d.) observations.

In the i.i.d. case, Rosenblatt(1956), observed that "all estimators of the density function satisfying relatively mild conditions are necessarily biased." Parzen (1962), gave asymptotic properties of kernel type estimators of a univariate density f on the basis of a random i.i.d sample $X_1, X_2, ..., X_n$ from f. Cacoullos (1964), extended these results to a p-variate density ($p \geq 2$).

Roussas (1969a), extended Parzen's results for the case of real-valued Markov chains satisfying a strong condition known as Doeblin's hypothesis and also proved corresponding results for the transition density. The same author (1991a), proved strong consistency for the estimator of the stationary density assuming that the chain is strictly stationary, ρ-mixing and satisfies some additional conditions. Rosenblatt (1970), considered the problem for stochastic processes satisfying geometric ergodicity. Prakasa Rao (1983), proposed a recursive estimator for f, and under Rosenblatt's condition proved convergence in quadratic mean. Roussas (1991b), also considered recursive estimators under ρ-mixing condition.

The goal of this dissertation is to relax some of the strong assumptions of Roussas and others in developing an inference theory for Markov Chains. The key assumptions of Roussas are:

(i) The chain $\{X_n$:n=0,1,2,...$\}$ satisfies Doeblin's condition,

(ii) The chain is already in steady state, i.e. X_0 has the stationary distribution,

(iii) The chain satisfies strong mixing decay conditions.

The second assumption rules out all initial distributions of X_0 other than the stationary one. In particular it rules out deterministic starting points.

There are a number of Markov chains that do not satisfy the Doeblin condition. For example take the autoregressive sequence $\{X_n :$ n=0,1,2,...$\}$ where

$X_{n+1} = \rho X_n + \epsilon_{n+1}$ with $|\rho| < 1$ and $\{\epsilon_n : n = 1, 2, ...\}$ i.i.d uniform $(-1, 1)$.

This Markov chain has stationary density f, but it is not Doeblin recurrent (see example 1.3.6-b below). So Roussas'results on the asymptotic properties of its estimator do not apply.

A more interesting example comes from problems of storage of water in dams or reservoirs. Let X_n be the content of the dam at time n. The development of $\{X_n$: n=0,1,2,...$\}$ is governed by the recursion

$$X_{n+1} = \min\{X_n + Y_{n+1}, K\} - \min\{X_n + Y_{n+1}, M\}$$

where K is the capacity of the dam, Y_n is the amount of water flowing to the dam during the time period $(n-1,n)$, and $M \in (0,K)$ is the amount of water which is released at time $n = 1,2,\ldots$. From a practical point of view, the stationary probability distribution π is of importance for assesing values of quantities like the proportion $\pi_o = \pi(0)$ of time the dam is empty. See Asmussen (1987), Moran (1959), and Stadje (1993).

The above model can be generalized, e.g. by taking M random instead of constant. These models are related to other models for storages and inventories if we think X_n as the stock level at time n, the release M_n corresponds to the amount being sold and the input Y_n to reordering of the material. See Prabhu (1980).

One of the problems faced in trying to establish asymptotic properties of kernel estimators is computing their variance. Doeblin condition implies that $\mathrm{Cov}(X_k, X_{k+r})$ decreases rapidly to zero as $r \to \infty$ and $k \to \infty$. See Doob (1953). This in turn, yields easily consistency and asymtotic normality of kernel type estimators of f.

A much weaker condition than Doeblin's was introduced by Harris (1956). That is the so called ϕ-recurrence or Harris recurrence. Athreya and Ney (1978) gave an equivalent definition of Harris recurrence. In their work they introduced the technique of regenerative processes for the study of the limit theory of a recurrent Harris chain. This was done independently by Nummelin (1978).

We will use the techniques of regenerative processes to propose estimators and prove their properties under the Harris recurrence condition and without requiring the two other conditions of stationarity and rapid mixing.

In the rest of this chapter we present basic definitions and the necessary results to our purpose. In the second chapter we discuss the naive kernel estimator. In the third chapter, the results of chapter 2 will be extended to a class of kernels satisfying some conditions. In the fourth chapter we present some results for semi- Markov processes.

1.2 Basic Definitions and Some Results

Definition 1.2.1 *Transition Function.* Let S be a nonempty set, Σ be a σ-algebra of subsets of S. A function $P(x, A){:}S \times \Sigma \to [0,1]$ is called a (Markov) transition function on (S, Σ) if

(*i*) $\forall x \in S$, $P(x, .)$ is a probability measure on Σ,

(*ii*) $\forall A \in \Sigma$, $P(., A)$ is a Σ-measurable function.

For $n \geq 0$ define the iterates $P^{(n)}$ of P by the recursive relation

$$P^{(n+1)}(x, A) = \int_S P(x, dy) P^{(n)}(y, A) \qquad (1.1)$$

where $P^{(1)} = P$ and $P^{(0)}(x, A) = 1$ if $x \in A$ and 0 if $x \in A^c$, i.e., the delta measure at x. Then for every n , $P^{(n)}(x, A)$ is also a Markov transition function and for all $m, n \geq 0$,

$$P^{(m+n)}(x, A) = \int_S P^{(m)}(x, dy) P^{(n)}(y, A) \qquad (1.2)$$

This is called the Chapman - Kolmogorov equation.

Definition 1.2.2 A sequence of S-valued random variables $\{X_n$:n=0,1,2...$\}$ is a *homogeneous Markov chain* with P as its transition function if and only if for every $n \geq 0$,

$$\begin{aligned} P(X_{n+1} \in A \mid \sigma(X_0, X_1, ..., X_n)) &= P(X_{n+1} \in A \mid \sigma(X_n)) \\ &= P(X_n, A) \end{aligned} \tag{1.3}$$

with probability 1, where for any collection D of r.v., $\sigma(D)$ is the σ- algebra generated by the r.v. in D and $P(. \mid \sigma(D))$ stands for conditional probability given the σ-algebra. We denote by P_μ, the probability measure of the sequence $\{X_n$:n=1,2.,...$\}$ when X_0 is distributed according to μ. We also write P_x for μ =delta measure at x. The distribution μ of X_0 is called the initial distribution of the Markov chain.

Definition 1.2.3 Let $\{X_n$:n=0,1,2...$\}$ be an S-valued stochastic process. A random variable τ is said to be a *stopping time* w.r.t. $\{X_n$:n=0,1,2...$\}$ if it assumes only nonnegative(integers) values and for every integer m, the event $\{\tau \leq m\}$ is $\sigma(X_0, X_1, ..., X_m)$-measurable. Typical examples of a stopping time is a hitting time $\tau_A = \inf\{n : n \geq 1, X_n \in A\}$ of a set $A \in \Sigma$.

Definition 1.2.4 *Strong Markov Property.* Let $\{X_n$:n=0,1,2...$\}$ be a stochastic process. We say that this process has the strong Markov property with respect to a stopping time τ if for every m,

$$P(X_{\tau+m} \in A \mid \Im_\tau) = P^m(X_\tau, A) \text{ w.p. } 1 \tag{1.4}$$

where $\Im_\tau$ is the σ-algebra $\{A : A \cap (\tau \leq m) \in \Im_m \forall m = 0, 1, 2, ...\}$.

Remark. The process is said to be strong Markov if it has the strong Markov property with respect to any stopping time τ.

Theorem 1.2.5 Any discrete-time Markov chain has the strong Markov property. This is corollary 6.2, chapter 1 in Asmussen (1987).

Definition 1.2.6 Let $\{X_n$:n=0,1,2...$\}$ be a Markov chain on (S, Σ) with transition function $P(x, A)$. π is a *stationary distribution* for the Markov Chain if for every $A \in \Sigma$,

$$\pi(A) \quad = \quad \int_S P(x, A)\pi(dx) \tag{1.5}$$

Thus if X_0 is distributed according to π, then so is X_1 and hence X_n for every $n \geq 1$. In this case the chain $\{X_n$:n=0,1,2,...$\}$ is a *strictly stationary sequence.*

1.3 Regeneration

When the state space is discrete, we know that a state j is recurrent if $P_j(\tau_j < \infty) = 1$ where τ_j is the hitting time of state j, defined as $\tau_j = \inf\{n : n \geq 1, X_n = j\}$. What is the corresponding definition when the state space is not discrete?. Note that if $\{X_n$:n=0,1,2,...$\}$ are continuous random variables then $P_x(X_n = x$ for some $n \geq 1) = 0$. Thus the process may not return to the starting point at all. Consider, again, for example the process

$X_{n+1} = \rho X_n + \epsilon_{n+1}$ with $|\rho| < 1$ and $\{\epsilon_n : n = 1, 2, ...\}$ i.i.d uniform $(-1, 1)$.

However the process is neighborhood recurrent in the sense $P_x(\tau_I < \infty) = 1$ for every open interval I and $x \in S = R$.

A very strong form of recurrence was given by

Condition 1.3.1 *Doeblin Condition.* Let $\{X_n$: n=0,1,2,...$\}$ be a Markov chain on (S, Σ). We say that the Doeblin condition is satisfied if there exist a probability measure ϕ on the state space (S, Σ), a real number $\epsilon > 0$, and an integer $n_0 > 0$ such that

$$\forall x \in S \text{ and } \forall A \in \Sigma,\ P^{(n_0)}(x, A) \geq \epsilon\phi(A).$$

In the general state space, this is a very strong condition. When S is finite and the chain is irreducible, Doeblin condition is always satisfied with $\phi(.)$ = delta measure of some singleton. However, if S is countable, even a chain that is irreducible and satisfies $P_j(\tau_j < \infty) = 1$ for every j, it need not to be Doeblin recurrent. See example 1.3.6-a below.

Definition 1.3.2 Let $\{X_n$:n=0,1,2,...$\}$ be a stochastic process. The maximal correlation of order k, denoted by $\rho(k)$, is defined as

$$\rho(k) \;=\; \sup\{|\rho(\xi,\eta)| : \xi \in \sigma(X_0, ..., X_r), \eta \in \sigma(X_{r+k}, ...), E|\xi|^2, E|\eta|^2 < \infty\}$$

where $\rho(\xi, \eta)$ is the correlation coefficient of ξ and η.

Definition 1.3.3 The sequence $\{X_n$:n=0,1,2,...$\}$ is called ρ-mixing if $\rho(k) \to 0$ as $k \to \infty$

Definition 1.3.4 A Markov chain $\{X_n$:n=0,1,2,...$\}$ is called *Harris recurrent* if there exists a σ-finite measure ϕ on the state space (S, Σ) such that $\phi(A) > 0$ implies

$$P_x(\tau_A < \infty) \;=\; 1\ \forall x \in S$$

where P_x denotes the probability starting at x. Any irreducible and recurrent countable state space Markov chain is Harris recurrent, but not necessarily Doeblin recurrent.

The following is an equivalent definition given by Athreya and Ney (1978)

Definition 1.3.5 A Markov chain $\{X_n$:n=0,1,2,...$\}$ is called (A, ϵ, ϕ, n_0) *recurrent* if there exist a set $A \in \Sigma$, a probability measure ϕ on A, a real number $\epsilon > 0$, and an integer $n_0 > 0$ such that

$$P_x(\tau_A < \infty) = P_x(X_n \in A \text{ for some } n \geq 1) = 1 \ \forall x \in S \tag{1.6}$$

$$P_x(X_{n_0} \in E) = P^{(n_0)}(x, E) \geq \epsilon\phi(E) \ \forall x \in A \text{ and } \forall E \subset A \tag{1.7}$$

Doeblin condition would say $A = S$. Definition 1.3.5 holds, e.g. with $A = \{\Delta\}$ for a singleton $\{\Delta\}$ if $P_x(X_n = \Delta \text{ for some } n \geq 1) = 1 \ \forall x \in S$. In this case take any $n_0 > 0$ such that $P_\Delta(X_{n_0} = \Delta) > 0$, $\epsilon = 0.5$ and $\phi(E) = P^{(n_0)}(\Delta, E)$.

In many practical cases, the second definition seems easier to check than the first one. For $S = R$, a natural choice of ϕ is Lebesgue measure (possibly restricted to some interval).

Example 1.3.6 Before going on, we give two examples of Harris recurrent chains that do not satisfy Doeblin condition.

a) Take the simple symmetric random walk, on the set of natural numbers, with reflexion at 0. Namely: $S = \{0, 1, 2, ...\}$ and the transition probabilities defined by $P(0,0) = P(0,1) = 0.5$, $P(i, i+1) = P(i, i-1) = 0.5$ for $i \geq 1$.

Let ϕ be any probability measure on S. Then there exists $i_0 \in S$ such that $\phi(i_0) > 0$. Now take $A = \{i_0\}$. Given n_0, $P^{(n_0)}(x, A) = 0$ for every $x > n_0 + i_0$. So Doeblin condition is not satisfied. Since the state 0 is recurrent and the chain is irreducible, $P_x(\tau_0 < \infty) = 1 \ \forall x \in S$. Thus, this chain is Harris recurrent.

b) Using the same idea we can prove that the autoregressive process considered earlier, on pp 6, is not Doeblin recurrent: Let ϕ be any probability measure on R. There exists an open interval $I = (a, b)$, $-\infty < a < b < \infty$, such that $\phi(I) > 0$. Given n_0, take x_0 such that $\rho^{n_0 - 1} x_0 < a - \frac{\rho}{1-\rho}$. Then $\lim_{x \to -\infty} P_x(X_{n_0} \in I) = 0$. Athreya and Pantula (1986) proved that this process is Harris recurrent. In fact, they proved that if $|\rho| < 1$ and $\sum_{j=1}^{n} \rho^j \epsilon_j$ does not have a purely singular distribution for some $n \geq 1$, then the autoregressive process turns out to be Harris recurrent.

The main tool to our purpose is the regeneration lemma proved by Athreya and Ney (1978)

Lemma 1.3.7 Regeneration Lemma If $\{X_n$:n=0,1,2,...$\}$ is (A, ϵ, ϕ, n_0) recurrent, then there exists a random time N such that $P_x(N < \infty) = 1$ and

$$\begin{aligned} a(x,n,k) &\stackrel{\text{def}}{=} P_x(X_n \in A, X_{n+1} \in A_1, ..., X_{n+k} \in A_k, N = n) \\ &= P_x(N = n) \int_A P_y(X_1 \in A_1, ..., X_k \in A_k)\phi(dy) \end{aligned}$$

That is, the evolution of the process for $n \geq N$ is independent of $X_1, X_2, ..., X_{N-1}$ and N and has the same distribution as $\{X_n$:n=0,1,2,...$\}$ where X_0 is distributed according to ϕ. Thus N is a random time such that the pre-N and post-N evolution are independent and the post-N process has a distribution independent of $X_0, ..., X_{N-1}$.

Corollary 1.3.8 If $\{X_n$:n=0,1,2,...$\}$ is (A, ϵ, ϕ, n_0) recurrent, then there exists a sequence of random times $N_i; i = 1, 2, ...$ such that X_{N_i} have distribution ϕ on A, and the random tours $\{X_{N_i+j} : j = 0, 1, 2, ..., N_{i+1} - N_i - 1; N_{i+1} - N_i\}$ are independent, identically distributed, and independent of N_1.

The regeneration lemma can be used to show the existence of a stationary measure for Harris recurrent chains.

Theorem 1.3.9 Define

$$\nu(E) = E_\phi \Sigma_{i=0}^{N_1 - 1} I(X_i \in E) \tag{1.8}$$

where N_1 is the first regeneration time as in Lemma 1.3.7.

Then ν is a stationary measure for $\{X_n$:n=0,1,2,...$\}$, and it is unique up to a multiplicative constant. Since $\nu(S) = E_\phi N_1$, ν is finite if and only if $E_\phi N_1 < \infty$.

Corollary 1.3.10 A stationary probability distribution $\pi(.)$ for $\{X_n\text{:n=0,1,2,...}\}$ exists if and only if $E_\phi N_1 < \infty$, and in this case,

$$\pi(E) \;=\; \frac{\nu(E)}{\nu(S)} \tag{1.9}$$

Proofs are in Athreya and Ney (1978).

Another characterization of the stationary measure ν is given by the following

Theorema 1.3.11 For a Harris recurrent chain, a measure ν satisfies (1.8) if and only if

$$\int_S g d\nu \;=\; E_\phi \Sigma_{i=0}^{N_1-1} g(X_i) \tag{1.10}$$

for every bounded measurable function $g : (S, \Sigma) \to R$.

Proof: Immediate from (1.8).

Theorem 1.3.12 Let g any bounded measurable function, then

$$E_\phi(\Sigma_{j=0}^{N_1-1} g(X_j))^2 \;=\; \int_S g^2(x)\nu(dx) + 2\int_S g(x)(Tg)(x)\nu(dx) \tag{1.11}$$

where $Tg(x) = E_x \Sigma_{j=1}^{N_1-1} g(X_j)$

Proof.

$$\begin{aligned} E_\phi(\Sigma_{j=0}^{N_1-1} g(X_j))^2 &= E_\phi \Sigma_{j=0}^{N_1-1} g^2(X_j) + 2E_\phi \Sigma_{j=0}^{N_1-1} \Sigma_{k=j+1}^{N_1-1} g(X_j)g(X_k) \\ &= \int_S g^2(x)\nu(dx) + 2E_\phi \Sigma_{j=0}^{N_1-1} \Sigma_{k=j+1}^{N_1-1} g(X_j)g(X_k) \end{aligned}$$

The second equality follows by Theorem 1.3.11. Now, by Markov Property,

$$\begin{aligned} E_\phi \Sigma_{j=0}^{N_1-1} \Sigma_{k=j+1}^{N_1-1} g(X_j)g(X_k) &= E_\phi \Sigma_{j=0}^{N1-1} g(X_j) E_{X_j} \Sigma_{k=1}^{N_1-1} g(X_k) \\ &= \int_S g(x)(E_x \Sigma_{k=1}^{N_1-1} g(X_k))\nu(dx) \\ &= \int_S g(x)(Tg)(x)\nu(dx) \end{aligned} \tag{1.12}$$

The second equality follows by Theorem 1.3.11 again.

1.4 Kernel Estimators

Our goal in the thesis is to develop appropriate estimators of the stationary distribution π and of the transition function P, and in particular for their densities with respect to the corresponding Lebesgue measure, when S is the real line.

Definition 1.4.1 Let $\{X_n\text{:n=0,1,2,...}\}$ be a homogeneous Markov chain with the real line as its state space and with transition function $P(x, A)$. Assume that there exists a stationary probability distribution π with density f with respect to Lebesgue measure. Assume also that the transition function admits density $t(y/x)$. Let K be a probability density function on R and $\{\delta_n\text{:n=0,1,2,...}\}$ be a sequence of positive numbers. If the chain is observed up to time n, define

$$f_n(x) = \frac{1}{n\delta_n}\Sigma_{i=0}^{n} K(\frac{x - X_i}{\delta_n}) \tag{1.13}$$

$$q_n(x,y) = \frac{1}{n\delta_n^2}\Sigma_{i=0}^{n-1} K(\frac{x - X_i}{\delta_n})K(\frac{y - X_{i+1}}{\delta_n}) \tag{1.14}$$

$$t_n(y \mid x) = \frac{q_n(x,y)}{f_n(x)} \tag{1.15}$$

In (1.15), if $f_n(x) = 0$, we define $t_n(y \mid x) = \delta_{xy}$. The random functions f_n and t_n are called kernel estimators of f and t respectively. We assume that K satisfies the following conditions:

$(K1)$ $K(.)$ is bounded, i.e., $\exists\ 0 < M < \infty \ni \forall x \in R,\ K(x) \leq M$

$(K2)$ $|x|K(x) \rightarrow 0$ as $|x| \rightarrow \infty$

$(K3)$ $\int_R x^2 K(x)dx < \infty$

Remark. Since K is a probability density, $(K1)$ implies $\int_R K^2(x)dx < \infty$.

As an example, take $K_0(.)$ be the uniform density on $(-1,1)$, that is, $K_0(x) =$

$\frac{1}{2}I_{(-1,1)}(x)$. Then, $f_n(.)$ reduces to

$$f_n(x) = \frac{1}{2\delta_n}\frac{1}{n}\sum_{j=0}^{n} I_{(x-\delta_n,x+\delta_n)}(X_j) = \frac{1}{2\delta_n}\frac{N_n(A_n)}{n} \tag{1.16}$$

where $N_n(A_n)$ is the number of visits to $A_n = (x-\delta_n, x+\delta_n)$ during $\{0,1,2,...,n\}$.

To finish this chapter we present, without proofs, two very useful results.

Theorem 1.4.2 Let (R^m, B_m) be the m-dimensional Euclidean space with the corresponding Borel σ-algebra , and let (R, B) be the real line.

Let $K : (R^m, B_m) \to (R, B)$ be measurable satisfying:

$$(i) \quad \exists\, 0 < M < \infty \ni |K(x)| \leq M \ \forall x \in R^m$$

$$(ii) \quad \int_{R^m} |K(x)|dx < \infty$$

$$(iii) \quad ||x|||K(x)| \to 0 \text{ as } ||x|| \to \infty$$

Furthermore, let $g : (R^m, B_m) \to (R, B)$ be measurable such that $\int_{R^m} |g(x)|dx < \infty$. Define

$$g_n(x) = \frac{1}{\delta_n}\int_{R^m} K(\frac{z}{\delta_n})g(x-z)dz$$

where $0 < \delta_n$, $\delta_n \to 0$ as $n \to \infty$. Then if g is continuous at x,

$$\lim_{n\to\infty} g_n(x) = g(x)\int_{R^m} K(z)dz$$

Proof: This is Theorem 2.1.1 in Prakasa Rao (1983). Note that K(.) need not be nonnegative.

Theorem 1.4.3 (Lebesgue density theorem) Let f be locally integrable in (R^m, B_m) with respect to Lebesgue measure. Then for almost all x

$$\frac{1}{(2\delta_n)^m}\int_{A_n} f(y)dy \to f(x)$$

where $A_n = \{y : x_i - \delta_n \leq y_i \leq x_i + \delta_n ; i = 1, 2, ..., m\}$ and $\delta_n \rightarrow 0$

Proof: See Rudin (1987) pp 138.

2. NAIVE KERNEL ESTIMATOR FOR REAL-VALUED HARRIS CHAINS

Let $\{X_n$: n=0,1,2,...$\}$ be a real-valued Markov chain. In this chapter we consider the naive kernel estimator for the case in which there is a recurrence point Δ. In the first section we will study estimator for the stationary density and in the second section we will study estimator for the transition density. Let T_Δ be the regeneration time, let $\lambda = E_\Delta T_\Delta$. Through this chapter we assume $E_\Delta T_\Delta^2 < \infty$. In section 2.3 we will show that all the results in sections 2.1 and 2.2 are valid for real-valued Harris recurrent chains by the use of the regeneration technique outlined in 1.1.3.

2.1 Kernel Estimator of the Stationary Density

Let $\pi(.)$ be the stationary probability distribution and assume that on $R \setminus \{\Delta\}$, $\pi(.)$ is an absolutely continuous distribution with density f with respect to Lebesgue measure.

The naive kernel estimator of f is defined as in (1.16)

$$p_n(x) \;\; = \;\; \frac{1}{2n\delta_n} \Sigma_{i=0}^{n} I(X_i \in A_n) \qquad (2.1)$$

where $A_n = (x - \delta_n, x + \delta_n)$. In this chapter we will denote the naive kernel estimator as $p_n(.)$ to distinguish it from the general case treated in chapter 3 where it will be denoted by $f_n(.)$.

2.1.1 Weak Consistency of p_n

The main result of this section is Theorem 2.1.4 below which establishes the weak consistency of $p_n(x)$ for $f(x)$. In what follows x will be a generic element in $R \setminus \{\Delta\}$. All conclusions asserted are supposed to hold for almost all x (w.r.t. Lebesgue measure) unless special assumptions such as continuity or differentiability of f at x are impossed. The first step is the following

Lemma 2.1.1 Define

$$\begin{aligned} T_\Delta^{(1)} &= \inf\{k > 0 : X_k = \Delta\} \\ T_\Delta^{(2)} &= \inf\{k > T_\Delta^{(1)} : X_k = \Delta\} \\ &\dots \\ T_\Delta^{(r)} &= \inf\{k > T_\Delta^{(r-1)} : X_k = \Delta\} \end{aligned}$$

For i=1,2, ... ; define

$$\eta_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} I(X_j \in A_n)$$

Then,

$$E_\Delta(\eta_{n1} - E_\Delta \eta_{n1})^2 = O(\delta_n) \tag{2.2}$$

Proof: Since $\{\eta_{ni} : i = 1, 2, ...\}$ are i.i.d., it is enough to prove the result for $i = 1$. By definition

$$\begin{aligned} (E_\Delta \eta_{n1})^2 &= (E_\Delta \Sigma_{j=0}^{T_\Delta^{(1)}-1} I(X_j \in A_n))^2 \\ &= (\lambda \int_R I_{A_n}(y) f(y) dy)^2 \qquad \text{(by Theorem 1.3.11)} \\ &= (\lambda \pi(A_n))^2 \\ &= O(\delta_n^2) \end{aligned}$$

Now,

$$\begin{aligned} E_\Delta \eta_{n1}^2 &= E_\Delta(\Sigma_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j))^2 \\ &= \lambda \int_R I_{A_n}(y) f(y) dy + 2\lambda \int_R I_{A_n}(y) E_y \Sigma_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) f(y) dy \\ &\quad \text{(by Theorem 1.3.12)} \end{aligned}$$

Now, $E_\Delta T_\Delta^2 < \infty$ implies $\int_R E_y(T_\Delta) f(y) dy < \infty$ and so, by the Lebesgue density theorem (Theorem 1.4.3), $\frac{1}{2\delta_n}\int_{A_n} f(y)dy$ and $\frac{1}{2\delta_n}\int_{A_n} E_y(T_\Delta) f(y) dy$ converge a.e. to $f(x)$ and $f(x)E_x(T_\Delta)$ respectively and for such x, bounded in n: the first integral on the right is equal to $\lambda\pi(A_n) = O(\delta_n)$ for almost all x. For the second integral, simply note that $E_y \Sigma_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) \leq E_y T_\Delta^{(1)}$ and so, the second integral is also $O(\delta_n)$ for almost all x. Thus $E_\Delta \eta_{n1}^2 = O(\delta_n)$. □

Lemma 2.1.2 Let $\{k_n : n = 1, 2, ...\}$ be a nonrandom sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. If $n\delta_n \to \infty$, then with η_{ni} as defined in Lemma 2.1.1,

$$\frac{1}{2k_n\delta_n}\Sigma_{i=1}^{k_n}\eta_{ni} \quad \to \quad \lambda f(x) \text{ in probability}$$

Proof:

$$\frac{1}{2k_n\delta_n}\Sigma_{i=1}^{k_n}\eta_{ni} \quad = \quad \frac{1}{2k_n\delta_n}\Sigma_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni}) + \frac{1}{2\delta_n}E_\Delta\eta_{n1}$$

Since $\frac{1}{2\delta_n}E_\Delta\eta_{n1} = \frac{\lambda\pi(A_n)}{2\delta_n} \to \lambda f(x)$, it is enough to prove that the first term converges to zero in probability. For any x for which (2.2) holds,

$$\begin{aligned} P_\Delta(\frac{1}{2k_n\delta_n}|\Sigma_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni})| > \epsilon) &\leq E_\Delta\left[\frac{1}{2\delta_n}\Sigma_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni})\right]^2 \\ &= \frac{k_n E_\Delta(\eta_{n1} - E_\Delta\eta_{n1})^2}{4\delta_n^2 k_n^2 \epsilon^2} \\ &= O(\frac{1}{k_n\delta_n}) \end{aligned}$$

Since $k_n\delta_n \to \infty$, the proof is complete. □

Lemma 2.1.3 Let $\{K_n$:n=1,2,...$\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p. 1, $0 < \alpha < \infty$. If $n\delta_n \to \infty$, then

$$\frac{1}{2K_n\delta_n}\Sigma_{i=1}^{K_n}\eta_{ni} \to \lambda f(x) \text{ in probability}$$

Proof: For $\epsilon > 0$, let $A(n,\alpha,\epsilon) = \{n(\alpha-\epsilon) < K_n < n(\alpha+\epsilon)\}$. For any $\theta > 0$

$$\begin{aligned} & P_\Delta\left[\frac{1}{2K_n\delta_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta\right] \\ = \; & P_\Delta\left[\frac{1}{2K_n\delta_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta, A(n,\alpha,\epsilon)\right] \\ & +P_\Delta\left[\frac{1}{2K_n\delta_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta, A^c(n,\alpha,\epsilon)\right] \end{aligned}$$

Since $\frac{K_n}{n} \to \alpha$ w. p. 1, the second term converges to zero. Now,

$$\begin{aligned} & P_\Delta\left[\frac{1}{2K_n\delta_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta, A(n,\alpha,\epsilon)\right] \\ \leq \; & P_\Delta\left[\frac{1}{[n(\alpha-\epsilon)]}\frac{1}{2\delta_n}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\eta_{ni} - \lambda f(x) > \theta, A(n,\alpha,\epsilon)\right] \\ \leq \; & P_\Delta\left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]}\frac{1}{[n(\alpha+\epsilon)]+1}\frac{1}{2\delta_n}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\eta_{ni} - \lambda f(x) > \theta\right] \end{aligned}$$

Since

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} = 1,$$

the last probability converges to zero by Lemma 2.1.2. Similarly it is proved that $P_\Delta(\frac{1}{2K_n\delta_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) < -\theta) \to 0$. □

Theorem 2.1.4 Let $\{X_n : n = 0,1,...\}$ be a real-valued Markov chain with stationary transition function. Let Δ be a real number such that for any $x \in R$, $P_x(T_\Delta < \infty) = 1$ where $T_\Delta = \inf\{n : n \geq 1, X_n = \Delta\}$. Assume $E_\Delta T_\Delta^2 < \infty$.

Let $\pi(A) = \frac{E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_A(X_j)}{E_\Delta T_\Delta^{(1)}}$ for any Borel set A in R. Assume that π is absolutely continuous w.r.t Lebesgue measure on $R \setminus \{\Delta\}$. Let f be the corresponding density. Let $\delta_n > 0$ and $p_n(x) = \frac{1}{2n\delta_n} \sum_{j=0}^{n} I_{A_n}(X_j)$ where $A_n = (x - \delta_n, x + \delta_n)$. If $\delta_n \to 0$ and $n\delta_n \to \infty$, then for almost all x and for every initial distribution,

$$p_n(x) \quad \to \quad f(x) \qquad \text{in probability}$$

Proof: We may decompose $p_n(.)$ as

$$\begin{aligned}
p_n(x) &= \frac{1}{2n\delta_n} \sum_{j=0}^{n} I(X_j \in A_n) \\
&= \frac{1}{2n\delta_n} \sum_{j=0}^{T_\Delta^{(1)}-1} I(X_j \in A_n) + \frac{1}{2n\delta_n} \sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} I(X_j \in A_n) \\
&\quad + \frac{1}{2n\delta_n} \sum_{j=T_\Delta^{(K_n)}}^{n} I(X_j \in A_n) \\
&= \alpha_n + \beta_n + \gamma_n \qquad \text{(say)}
\end{aligned}$$

where K_n is the random number of cycles, i.e., the number of visits to Δ during $\{0, 1, 2, ...\}$. Since $n\delta_n \to \infty$, $\alpha_n \to 0$ w.p.1. It is known that $E_\Delta T_\Delta^{(1)} < \infty$ implies that the family $\{n - T_\Delta^{(K_n)} : n = 1, 2, ...\}$ is tight (See Lemma 2.1.8 below). Then $\frac{n - T_\Delta^{(K_n)}}{n\delta_n} \to 0$ in probability. Hence, $\gamma_n \to 0$ in probability.

$$\begin{aligned}
\beta_n &= \frac{1}{2n\delta_n} \sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} I(X_j \in A_n) \\
&= \frac{1}{2n\delta_n} \sum_{i=1}^{K_n} \eta_{ni} \\
&= \frac{K_n}{n} \frac{1}{K_n} \sum_{i=1}^{K_n} \eta_{ni}
\end{aligned}$$

Now the result follows from Lemma 2.1.3 and the fact that $\frac{K_n}{n} \to \lambda^{-1}$ w.p. 1. $\square$

2.1.2 Asymptotic Normality of p_n

Next, we will prove asymptotic normality of p_n. The main results are Theorem 2.1.9 and Theorem 2.1.10 below. First consider the following

Lemma 2.1.5 Let the transition function $P(x,.)$ admit for all x a density $t(. \mid x)$ w.r.t. Lebesgue measure. Then,

$$\frac{E_\Delta(\eta_{n1} - E_\Delta\eta_{n1})^2}{\pi(A_n)} \longrightarrow \lambda \text{ as } n\delta_n \longrightarrow \infty$$

Proof: From Lemma 2.1.1

$$\begin{aligned}
E_\Delta(\eta_{n1} - E_\Delta\eta_{n1})^2 &= \lambda\pi(A_n) + 2\lambda\int_R I_{A_n}(y)E_y\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)f(y)dy \\
&\quad +o(\delta_n) \\
\int_{A_n} E_y\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)f(y)dy &= \int_{A_n}\left[E_y\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) - E_x\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)\right]f(y)dy \\
&\quad +(E_x\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j))\pi(A_n) \\
&= \alpha_n + \beta_n \text{ (say)}
\end{aligned}$$

$$\begin{aligned}
\alpha_n &= \int_{A_n}\left[E_y\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) - E_x\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)\right]f(y)dy \\
&= \int_{A_n}\left[E_y(\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) : T_\Delta^{(1)} \le k) - E_x(\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) : T_\Delta^{(1)} \le k)\right]f(y)dy \\
&\quad +\int_{A_n}\left[E_y(\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) : T_\Delta^{(1)} > k) - E_x(\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) : T_\Delta^{(1)} > k)\right]f(y)dy \\
&= \alpha_{n1} + \alpha_{n2} \text{ (say)}
\end{aligned}$$

Since $E_x(T_\Delta^{(1)} : T_\Delta^{(1)} > k) \to 0$ as $k \to \infty$; for $\epsilon > 0$, there exists k_0 such that $E_x(T_\Delta^{(1)} : T_\Delta^{(1)} > k) < \epsilon \ \forall k \geq k_0$. Also α_{n2} is bounded above by

$$\begin{aligned} & \int_{A_n} E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > k) f(y)dy + \int_{A_n} E_x(T_\Delta^{(1)} : T_\Delta^{(1)} > k) f(y)dy \\ = \ & \alpha_{n21} + \alpha_{n22} \qquad\qquad \text{(say)} \end{aligned}$$

By the Lebesgue density theorem, both $\frac{\alpha_{n21}}{2\delta_n}$ and $\frac{\alpha_{n22}}{2\delta_n}$ converge to $E_x(T_\Delta^{(1)} : T_\Delta^{(1)} > k)f(x)$. So $\overline{\lim}_n \frac{\alpha_{n2}}{2\delta_n} \leq \epsilon$. Now, for each $0 \leq r \leq k_0$, $0 \leq j \leq r$,

$$\begin{aligned} \int_{A_n} E_y\left[I_{A_n}(X_j) : T_\Delta^{(1)} = r\right] f(y)dy &\leq \int_{A_n} P_y(X_j \in A_n) f(y)dy \\ &= \int_{A_n}\int_{A_n} t^{(j)}(x \mid y)dx f(y)dy \end{aligned}$$

By the Lebesgue density theorem, the last integral is $O(\delta_n^2)$. Since α_{n1} is a sum of a finite number of terms, $\frac{\alpha_{n1}}{\pi(A_n)}$ is $O(\delta_n)$. Clearly $\frac{\beta_n}{\pi(A_n)}$ converges to zero and this completes the proof of the lemma. $\square$

Lemma 2.1.6 Let $\{k_n\text{:n=1,2,...}\}$ be a sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. Define

$$Z_n = \frac{1}{\sqrt{2\lambda k_n \delta_n f(x)}} \sum_{i=1}^{k_n} (\eta_{ni} - \lambda\pi(A_n))$$

If $n\delta_n \to \infty$ and $\delta_n \to 0$, then $Z_n \xrightarrow{d} N(0,1)$

Proof:

$$\begin{aligned} E_\Delta(\eta_{ni} - \lambda\pi(A_n)) &= 0 \\ \mathrm{Var}_\Delta(Z_n) &= \frac{1}{2\lambda k_n \delta_n f(x)} k_n E_\Delta(\eta_{ni} - \lambda\pi(A_n))^2 \to 1 \\ & \quad \text{(by Lemma 2.1.5)} \end{aligned}$$

So, it suffices to check Lindeberg's condition: Let us write

$$Z_n = \frac{1}{\sqrt{2\lambda k_n \delta_n f(x)}} \sum_{i=1}^{k_n} (\eta_{ni} - \lambda\pi(A_n))$$

$$= \sum_{i=1}^{k_n} W_{ni}, \text{ (say)}$$

Then,

$$\begin{aligned}
L_n(\epsilon) &= \sum_{i=1}^{k_n} E_\Delta(W_{ni}^2 : |W_{ni}| > \epsilon) \\
&= k_n E_\Delta(W_{n1}^2 : |W_{n1} > \epsilon) \\
&= \frac{k_n}{2\lambda k_n \delta_n f(x)} E_\Delta\left[(\eta_{n1} - \lambda\pi(A_n))^2 : |\eta_{ni} - \lambda\pi(A_n)| > \epsilon\sqrt{2\lambda k_n \delta_n f(x)}\right] \\
&= \frac{2}{2\lambda\delta_n f(x)} E_\Delta\left[\eta_{n1}^2 : |\eta_{n1} - \lambda\pi(A_n)| > \epsilon\sqrt{2\lambda k_n \delta_n f(x)}\right] \\
&\quad + \frac{1}{\delta_n f(x)} \lambda\pi^2(A_n) P_\Delta\left[|\eta_{n1} - \lambda\pi(A_n)| > \epsilon\sqrt{2\lambda k_n \delta_n f(x)}\right]
\end{aligned}$$

Since $\frac{\pi(A_n)}{2\delta_n} \rightarrow f(x)$, the second term converges to zero. Now, since $\eta_{n1} \geq 0$ w.p.1, $\frac{1}{\delta_n} E_\Delta\left[\eta_{n1}^2 : |\eta_{n1} - \lambda\pi(A_n)| > \epsilon\sqrt{2\lambda k_n \delta_n f(x)}\right]$ is bounded above by

$$\begin{aligned}
c_n &\stackrel{\text{def}}{=} \frac{1}{\delta_n} E_\Delta\left[\eta_{n1}^2 : \eta_{n1} > \epsilon' \sqrt{2\lambda k_n \delta_n f(x)}\right] \qquad \text{(where } \epsilon' > \epsilon) \\
&= \frac{1}{\delta_n} E_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) : \eta_{n1} > \epsilon' \sqrt{2\lambda k_n \delta_n f(x)}\right] \\
&\quad + \frac{2}{\delta_n} E_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{A_n}(X_k) : \eta_{n1} > \epsilon' \sqrt{2\lambda k_n \delta_n f(x)}\right] \\
&= c_{n1} + c_{n2}, \text{ (say)}
\end{aligned}$$

By the computations in Lemma 2.1.5, c_{n2} converges to zero. Also c_{n1} is bounded by

$$\begin{aligned}
c_{n1}' &\stackrel{\text{def}}{=} \frac{1}{\delta_n}\left[E_\Delta \eta_{n1}^2 P_\Delta(\eta_{n1} > \epsilon' \sqrt{2\lambda k_n \delta_n f(x)}\right]^{\frac{1}{2}} \\
&\leq \frac{1}{\delta_n}\left[\frac{(E_\Delta \eta_{n1}^2)^2}{2(\epsilon')^2 \lambda k_n \delta_n f(x)}\right]^{\frac{1}{2}} \quad \text{(by Schwartz and Chebychev's inequalities)} \\
&= O(\frac{1}{k_n \delta_n})^{\frac{1}{2}}
\end{aligned}$$

which converges to zero since $k_n\delta_n \to \infty$. Hence $L_n(\epsilon) \to 0$ and the proof of the lemma is complete. $\square$

Lemma 2.1.7 The conclusion of Lemma 2.1.6 holds if $\{K_n$:n=1,2,...$\}$ is a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p. 1 $0 < \alpha < \infty$.

Proof: The proof is similar to the proof of Lemma 2.1.3.

To finish this section we will prove asymptotic normality of p_n and if $x \neq y$ we will prove that $\sqrt{n\delta_n}p_n(x)$ and $\sqrt{n\delta_n}p_n(y)$ are asymptotically independent.

Before doing that we present, without proving, the following result from Renewal Theory.

Lemma 2.1.8 Let $\{X_n$:n=0,1,...$\}$ be a Harris chain with a recurrence point Δ. Let $\lambda = E_\Delta T_\Delta$ and $\sigma^2 = \mathrm{Var}_\Delta(T_\Delta)$. Let K_n be the random number of visits to Δ by the chain during $\{0, 1, 2, ...\}$. That is $K_n = \sum_{j=0}^n I(X_j = \Delta)$. Then the family $\{n - T_\Delta^{(K_n)} : n = 1, 2, ...\}$ is tight and

$$\frac{K_n}{n} \to \frac{1}{\lambda} \text{ a.e. as } n \to \infty \tag{2.3}$$

$$\sqrt{n}\left(\frac{K_n}{n} - \frac{1}{\lambda}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2}{\lambda^3}\right) \tag{2.4}$$

Assertions (2.3) and (2.4) follow from the discrete version of proposition 1.4, chapter IV and proposition 4.3, chapter VI in Asmussen (1987). The first assertion is also derivable from the same material.

Theorem 2.1.9 If $\delta_n \to 0$, $n\delta_n \to \infty$ as $n \to \infty$, then

$$\sqrt{n\delta_n}\left[p_n(x) - \frac{\pi(A_n)}{2\delta_n}\right] \xrightarrow{d} N\left(0, \frac{1}{2}f(x)\right)$$

Proof:

$$\sqrt{n\delta_n}\left[p_n(x) - \frac{\pi(A_n)}{2\delta_n}\right] = \sqrt{n\delta_n}\left[\frac{1}{2n\delta_n}\sum_{i=1}^{K_n}\eta_{ni} - \frac{\pi(A_n)}{2\delta_n}\right]$$

$$
\begin{aligned}
& + \frac{1}{2\sqrt{n\delta_n}} \left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) + \sum_{j=T_\Delta^{(k_n)}}^{n} I_{A_n}(X_j) \right] \\
= \; & W_n + W_n' \qquad \text{(say)}
\end{aligned}
$$

Since $n\delta_n \to \infty$ and the family $\{n - T_\Delta^{(K_n)} : n = 1, 2, ..\}$ is tight, $\frac{n - T_\Delta^{(K_n)}}{\sqrt{n\delta_n}} \to 0$ in probability. So it is enough to show that

$$
W_n = \sqrt{n\delta_n} \left[\frac{1}{2n\delta_n} \sum_{i=1}^{K_n} \eta_{ni} - \frac{\pi(A_n)}{2\delta_n} \right] \xrightarrow{d} N(0, \frac{1}{2} f(x))
$$

From Lemma 2.1.7,

$$
T_n \stackrel{\text{def}}{=} \frac{1}{2\sqrt{\lambda k_n \delta_n}} \sum_{i=1}^{K_n} (\eta_{ni} - \lambda \pi(A_n)) \xrightarrow{d} N(0, \frac{1}{2} f(x))
$$

Since $\frac{K_n}{n} \to \lambda^{-1}$ w.p. 1 we conclude that

$$
T_n' \stackrel{\text{def}}{=} \frac{1}{2\sqrt{n\delta_n}} \sum_{i=1}^{K_n} (\eta_{ni} - \lambda \pi(A_n)) \xrightarrow{d} N(0, \frac{1}{2} f(x))
$$

But

$$
\begin{aligned}
T_n' &= \sqrt{n\delta_n} \left[\frac{1}{2n\delta_n} \sum_{i=1}^{K_n} \eta_{ni} - \frac{K_n \lambda \pi(A_n)}{2n\delta_n} \right] \\
&= \sqrt{n\delta_n} \left[\frac{1}{2n\delta_n} \sum_{i=1}^{K_n} \eta_{ni} - \frac{\pi(A_n)}{2\delta_n} \right] \\
&\quad + \frac{\pi(A_n)}{2\delta_n} \lambda \sqrt{\delta_n} \sqrt{n} (\frac{1}{\lambda} - \frac{K_n}{n}) \\
&= W_n + W_n'' \text{ (say)}
\end{aligned}
$$

Since $\frac{\pi(A_n)}{2\delta_n} \to f(x)$ and $\delta_n \to 0$, $W_n'' \to 0$ in probability by Lemma 2.1.8. So $W_n \xrightarrow{d} N(0, \frac{1}{2} f(x))$. □

Now, observe that

$$
\sqrt{n\delta_n}(p_n(x) - f(x)) = \sqrt{n\delta_n}(p_n(x) - \frac{\pi(A_n)}{2\delta_n}) + \sqrt{n\delta_n}(\frac{\pi(A_n)}{2\delta_n} - f(x))
$$

We will find conditions under which the second term converges to zero. Suppose that f is twice differentiable at x. Then

$$\begin{aligned}\frac{\pi(A_n)}{2\delta_n} - f(x) &= \frac{1}{2\delta_n}\int_{A_n}(f(u)-f(x))du \\ &= \frac{1}{2\delta_n}\int_{A_n}\left[(u-x)f'(x)+\frac{1}{2}(u-x)^2f''(x)+o(\delta_n^2)\right]du \\ &= \frac{1}{6}f''(x)\delta_n^2+o(\delta_n^2)\end{aligned}$$

So we have the following

Theorem 2.1.10 If $\delta_n \to 0$, $n\delta_n \to \infty$, $n\delta_n^p \to 0$ for some $1 < p \leq 5$ and f is twice differntiable at x, then

$$\sqrt{n\delta_n}(p_n(x)-f(x)) \xrightarrow{d} N(0,\frac{1}{2}f(x)).$$

Finally, consider $x \neq y$, x and $y \neq \Delta$, and define

$$p_n(y) = \frac{1}{2n\delta_n}\sum_{j=0}^{n} I_{B_n}(X_j)$$

where $B_n = (y-\delta_n, y+\delta_n)$

Theorem 2.1.11 If $\delta_n \to 0$, $n\delta_n \to \infty$, then $\sqrt{n\delta_n}p_n(x)$ and $\sqrt{n\delta_n}p_n(y)$ are asymptotically independent.

Proof: For $i = 1, 2, ...$, consider η_{ni} as defined earlier and define

$$\tau_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} I_{B_n}(X_j)$$

For any l_1, l_2,

$$\begin{aligned}& l_1\sqrt{n\delta_n}\left[p_n(x)-\frac{\pi(A_n)}{2\delta_n}\right]+l_2\sqrt{n\delta_n}\left[p_n(y)-\frac{\pi(B_n)}{2\delta_n}\right] \\ =\;& l_1\sqrt{n\delta_n}\left[\frac{1}{n\delta_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} I(X_j\in A_n)-\frac{\pi(A_n)}{2\delta_n}\right]\end{aligned}$$

$$+l_2\sqrt{n\delta_n}\left[\frac{1}{n\delta_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} I(X_j \in B_n) - \frac{\pi(B_n)}{2\delta_n}\right]$$

$$+l_1\frac{1}{\sqrt{n\delta_n}}\sum_{j=0}^{T_\Delta^{(1)}-1} I(X_j \in A_n) + l_2\frac{1}{\sqrt{n\delta_n}}\sum_{j=0}^{T_\Delta^{(1)}-1} I(X_j \in B_n)$$

$$+l_1\frac{1}{\sqrt{n\delta_n}}\sum_{j=T_\Delta^{(K_n)}}^{n} I(X_j \in A_n) + l_2\frac{1}{\sqrt{n\delta_n}}\sum_{j=T_\Delta^{(K_n)}}^{n} I(X_j \in B_n)$$

and proceeding as in the proof of Theorem 2.1.9, we can show that the above converges to a normal distribution with mean zero and variance given by

$$\frac{1}{2}(l_1^2 f(x) + l_2^2 f(y)) + 2l_1 l_2 \sigma_{12}$$

where $\sigma_{12} = lim_{n\to\infty}\mathrm{Cov}_\Delta(\eta_{n1}, \tau_{n1})$

We now show that $\sigma_{12} = 0$:

$$\begin{aligned}
\mathrm{Cov}_\Delta(\eta_{n1}, \tau_{n1}) &= \mathrm{Cov}_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j), \sum_{j=0}^{T_\Delta^{(1)}-1} I_{B_n}(X_j)\right] \\
&= E_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) \sum_{j=0}^{T_\Delta^{(1)}-1} I_{B_n}(X_j)\right] - \lambda^2\pi(A_n)\pi(B_n) \\
&= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_j) + E_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) \sum_{k=j+1}^{T_\Delta^{(1)}-1} I_{B_n}(X_k)\right] \\
&\quad + E_\Delta\left[\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) \sum_{k=1}^{j-1} I_{B_n}(X_k)\right] - \lambda^2\pi(A_n)\pi(B_n)
\end{aligned}$$

Now,

$$\frac{1}{\delta_n}E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_j) = \frac{\lambda}{\delta_n}\int I_{A_n}(u) I_{B_n}(u) f(u)du$$

Since $x \neq y$, $A_n \cap B_n = \emptyset$ for n large enough. So there exists n_0 such that this integral is equal to zero for $n \geq n_0$. Also,

$$\begin{aligned}
\frac{1}{\delta_n} E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) \sum_{k=j+1}^{T_\Delta^{(1)}-1} I_{B_n}(X_k) \right] &= \frac{1}{\delta_n} E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) E_{X_j} \sum_{k=1}^{T_\Delta^{(1)}-1} I_{B_n}(X_k) \right] \cdot \\
&= \frac{\lambda}{\delta_n} \int I_{A_n}(u) E_u \left[\sum_{k=1}^{T_\Delta^{(1)}-1} I_{B_n}(X_k) \right] f(u) du \\
&= \frac{\lambda}{\delta_n} \int_{A_n} E_u \left[\sum_{k=1}^{T_\Delta^{(1)}-1} I_{B_n}(X_k) \right] f(u) du
\end{aligned}$$

It was proved in Lemma 2.1.5 that this expression converges to zero. Similarly it is proved that $\frac{1}{\delta_n} E_\Delta (\sum_{j=1}^{T_\Delta^{(1)}-1} \sum_{k=0}^{j-1} I_{B_n}(X_k))$ converges to zero. This implies that $\mathrm{Cov}_\Delta(\eta_{n1}, \tau_{n1}) = o(\delta_n)$. □

Corollary 2.1.12 Let f be twice differentiable at $x_1, x_2, ..., x_k$ in $R \setminus \{\Delta\}$. Let $\delta_n \to 0$, $n\delta_n \to \infty$, and $n\delta_n^p \to 0$ for some $1 < p \leq 5$. Then

$$\sqrt{2n\delta_n}\left(\frac{p_n(x_i) - f(x_i)}{\sqrt{p_n(x_i)}}, i = 1, 2, ..., k\right) \xrightarrow{d} (\xi_1, \xi_2, ..., \xi_k),$$

where $\xi_1, \xi_2, ..., \xi_k$ are independent and identically distibuted with $\xi_1 \sim N(0,1)$.

2.2 Kernel Estimator of the Transition Density

In this section we will prove consistency and asymptotic normality of q_n and t_n as defined in 1.4.1 with K, the naive kernel. In this case

$$q_n(x,y) = \frac{1}{4n\delta_n^2} \sum_{j=0}^{n-1} I_{A_n \times B_n}((X_j, X_{j+1}))$$

where $A_n = (x - \delta_n, x + \delta_n)$ and $B_n = (y - \delta_n, y + \delta_n)$; x a continuity point of $f(.)$ and $f(x) > 0$. In what follows, all assertions are supposed to hold for almost all (x, y) w.r.t. Lebesgue measure in R^2 except when specific smoothness assumptions of f and q are made at particular points.

2.2.1 Weak Consistency of q_n and t_n

The main results of the first part of this section are Theorem 2.2.4 and Theorem 2.2.5 below which establish the weak consistency of q_n and t_n respectively.

Lemma 2.2.1 For i=1,2,...Define

$$\Psi_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} I_{A_n \times B_n}((X_j, X_{j+1}))$$

Then

$$E_\Delta(\Psi_{ni} - E_\Delta \Psi_{ni})^2 = O(\delta_n^2)$$

Proof:

$$\begin{aligned} E_\Delta \Psi_{n1} &= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_{j+1}) \\ &= \lambda \int\int I_{A_n}(u) I_{B_n}(v) q(u,v)\,du\,dv \\ &= \lambda\gamma(A_n \times B_n) \\ &= O(\delta_n^2) \end{aligned}$$

where γ is the stationary distribution of $\{(X_n, X_{n+1}) : n = 0, 1, 2, ...\}$

$$\begin{aligned} E_\Delta \Psi_{n1}^2 &= E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_{j+1}) \right]^2 \\ &= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_{j+1}) \\ &\quad +2E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_{j+1}) I_{A_n}(X_k) I_{B_n}(X_{k+1}) \\ &= a_n + b_n \text{ (say)} \end{aligned}$$

$$
\begin{aligned}
b_n &= 2E_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)I_{B_n}(X_{j+1})E_{X_{j+1}}\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1})\right] \\
&= 2\lambda\int_{A_n}\int_{B_n} E_v\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1})q(u,v)dudv \\
&\le 2\lambda\int_{A_n}\int_{B_n} E_v(T_\Delta^{(1)})q(u,v)dudv \\
&= 2\lambda(4\delta_n^2)\int_{A_n}\frac{1}{2\delta_n}\int_{B_n}\frac{1}{2\delta_n}E_v(T_\Delta^{(1)})q(u,v)dudv
\end{aligned}
$$

By the Lebesgue density theorem, the last integral converges to $q(x,y)E_y(T_\Delta^{(1)}) < \infty$ and so $b_n = O(\delta_n^2)$. Also by the same theorem $\frac{a_n}{\delta_n^2} \to \lambda q(x,y)$. □

Lemma 2.2.2 Let $\{k_n$:n=1,2,...$\}$ be a nonrandom sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. Then

$$\frac{1}{4k_n\delta_n^2}\sum_{i=1}^{k_n}\Psi_{ni} \quad\to\quad \lambda q(x,y) \text{ in probability}$$

provided $\delta_n \to 0$, $n\delta_n^2 \to \infty$.

Proof: The proof is similar to the proof of Lemma 2.1.2 in this chapter.

Lemma 2.2.3 Lemma 2.2.2 holds if $\{K_n$:n=1,2,...$\}$ is a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p. 1, $0 < \alpha < \infty$.

Proof: Let $\epsilon > 0$, $\theta > 0$ and define $A(n,\alpha,\epsilon) = \{n(\alpha-\epsilon) < K_n < n(\alpha+\epsilon)\}$

$$
\begin{aligned}
& P_\Delta\left(\frac{1}{4K_n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni} - \lambda q(x,y) > \theta\right) \\
&= P_\Delta\left(\frac{1}{4K_n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni} - \lambda q(x,y) > \theta, A(n,\alpha,\epsilon)\right) \\
&\quad + P_\Delta\left(\frac{1}{4K_n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni} - \lambda q(x,y) > \theta, A^c(n,\alpha,\epsilon)\right) \\
&= \alpha_{n1} + \alpha_{n2} \qquad\qquad \text{(say)}
\end{aligned}
$$

Since $\frac{K_n}{n} \to \alpha$ w.p. 1, $P_\Delta(A^c(n,\alpha,\epsilon)$ and hence α_{n2} converges to zero. Now, since $\Psi_{ni} \geq 0$,

$$\begin{aligned}\alpha_{n1} &\leq P_\Delta\left[\frac{1}{4\left[n(\alpha-\epsilon)\right]\delta_n^2}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\Psi_{ni} - \lambda q(x,y) > \theta, A(n,\alpha,\epsilon)\right] \\ &\leq P_\Delta\left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]}\frac{1}{[n(\alpha+\epsilon)]+1}\frac{1}{4\delta_n^2}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\Psi_{ni} - \lambda q(x,y) > \theta\right]\end{aligned}$$

Since $\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} = 1$, the last probability converges to zero by Lemma 2.2.2. Similarly it is proved that $P_\Delta(\frac{1}{4K_n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni} - \lambda q(x,y) < -\theta) \to 0$. □

Theorem 2.2.4

$$q_n(x,y) \quad\to\quad q(x,y) \text{ in probability}$$

provided $\delta_n \to 0$, $n\delta_n^2 \to \infty$.

Proof:

$$\begin{aligned}q_n(x,y) &= \frac{1}{4n\delta_n^2}\sum_{j=0}^{n-1} I_{A_n}(X_j)I_{B_n}(X_{j+1}) \\ &= \frac{1}{4n\delta_n^2}\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)I_{B_n}(X_{j+1}) + \frac{1}{4n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni} \\ &\quad+\frac{1}{n\delta_n^2}\sum_{j=T_\Delta^{(K_n)}}^{n-1} I_{A_n}(X_j)I_{B_n}(X_{j+1})\end{aligned}$$

The first sum converges to zero w.p. 1 because $n\delta_n^2 \to \infty$ and $P_\Delta(T_\Delta^{(1)} < \infty) = 1$. Since $\frac{n-T_\Delta^{(K_n)}}{n\delta_n^2}$ converges to zero in probability, by Lemma 2.1.8, the last sum converges to zero in probability. Now, observe that

$$\frac{1}{n\delta_n^2}\sum_{i=1}^{4K_n}\Psi_{ni} \quad=\quad \frac{K_n}{n}\frac{1}{4K_n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni}$$

By Lemma 2.2.3 and Lemma 2.1.8, the last expression converges to $q(x,y)$ in probability.

To finish the first part of this section, by using Theorem 2.1.4 and Theorem 2.2.4, we establish the weak consistency of the estimator of the transition density.

Theorem 2.2.5 If $\delta_n \to 0$, $n\delta_n^2 \to \infty$, and $f(x) > 0$,

$$t_n(y \mid x) \quad \longrightarrow \quad t(y \mid x) \text{ in probability}$$

where t_n is as in (1.1.15).

2.2.2 Asymptotic Normality of q_n and t_n

To finish this section, we will prove asymptotic normality of $q_n(x,y)$ and $t_n(y \mid x)$.

Lemma 2.2.6 If $E_\Delta T_\Delta^2 < \infty$ and $\delta_n \to 0$, then

$$\frac{1}{4\delta_n^2} E_\Delta(\Psi_{ni} - E_\Delta \Psi_{ni})^2 \quad \longrightarrow \quad \lambda q(x,y)$$

Proof: By the computations in the proof of Lemma 2.2.1,

$$\begin{aligned}
\frac{1}{4\delta_n^2} E_\Delta(\Psi_{ni} - E_\Delta \Psi_{ni})^2 &= \frac{1}{4\delta_n^2} E_\Delta \Psi_{n1}^2 - \frac{1}{4\delta_n^2}(E_\Delta \Psi_{n1})^2 \\
&= \lambda \int_{A_n} \frac{1}{2\delta_n} \int_{B_n} \frac{1}{2\delta_n} q(u,v)\,du\,dv \\
&\quad + 2\lambda \int_{A_n} \int_{B_n} \frac{1}{4\delta_n^2} E_v \sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k) I_{B_n}(X_{k+1}) q(u,v)\,du\,dv \\
&\quad - \frac{1}{4\delta_n^2}(E_\Delta \Psi_{n1})^2
\end{aligned}$$

The third term is $O(\delta_n^2)$. By the Lebesgue density theorem, the first integral converges to $\lambda q(x,y)$. So it remains to prove that the second integral converges to zero. This integral is equal to

$$I_n \;=\; \int_{A_n} \frac{1}{2\delta_n} \int_{B_n} \frac{1}{2\delta_n} E_v \sum_{k=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_k) I_{B_n}(X_{k+1}) q(u,v)\,du\,dv$$

Now, observe that

$$
\begin{aligned}
E_v \sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}) &= E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}), T_\Delta^{(1)} > j\right] \\
&\quad + E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}), T_\Delta^{(1)} \le j\right] \\
&\le E_v(T_\Delta^{(1)} : T_\Delta^{(1)} > j) \\
&\quad + \sum_{r=1}^{j} E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}) : T_\Delta^{(1)} = r\right]
\end{aligned}
$$

Then,

$$
\begin{aligned}
I_n &\le \int_{A_n} \frac{1}{2\delta_n}\int_{B_n}\frac{1}{2\delta_n} E_v(T_\Delta^{(1)} : T_\Delta^{(1)} > j)q(u,v)dudv \\
&\quad + \int_{A_n} \frac{1}{2\delta_n}\int_{B_n}\frac{1}{2\delta_n}\sum_{r=1}^{j} E_v\left[\sum_{k=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}) : T_\Delta^{(1)} = r\right] q(u,v)dudv \\
&= I_{n1} + I_{n2} \text{ (say)}
\end{aligned}
$$

By Lebesgue density Theorem, I_{n1} converges to $E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > j)q(x,y)$. Then

$$
\begin{aligned}
&\overline{\lim}_{n\to\infty} \int_{A_n}\int_{B_n} \frac{1}{4\delta_n^2} E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}) : T_\Delta^{(1)} > j\right] q(u,v)dudv \\
\le\ & E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > j)q(x,y)
\end{aligned}
$$

Since $\lim_{j\to\infty} E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > j)q(x,y) = 0$, we conclude that given $\epsilon > 0$, $\exists\ j(\epsilon)$ such that for $j > j(\epsilon)$

$$
\begin{aligned}
&\overline{\lim}_{n\to\infty} \int_{A_n}\int_{B_n} \frac{1}{4\delta_n^2} E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_k)I_{B_n}(X_{k+1}) : T_\Delta^{(1)} > j\right] q(u,v)dudv \\
<\ & \epsilon
\end{aligned}
$$

Now, for each $0 \le r \le j(\epsilon)$, $0 \le k \le r$,

$$
\int_{A_n}\int_{B_n} \frac{1}{4\delta_n^2} E_v\left[I_{A_n}(X_k)I_{B_n}(X_{k+1}) : T_\Delta^{(1)} = r\right] q(u,v)dudv
$$

$$
\begin{aligned}
&\leq \int_{A_n}\int_{B_n} \frac{1}{4\delta_n^2} E_v\left[I_{A_n}(X_k)I_{B_n}(X_{k+1})\right] q(u,v)dudv \\
&= \int_{A_n}\int_{B_n} \frac{1}{4\delta_n^2}\left[\int\int I_{A_n}(z)I_{B_n}(w)t^{(k)}(z \mid v)t(w \mid z)dzdw\right] q(u,v)dudv \\
&= 4\delta_n^2 \int_{A_n}\frac{1}{2\delta_n}\int_{B_n}\frac{1}{2\delta_n}\int_{A_n}\frac{1}{2\delta_n}\int_{B_n}\frac{1}{2\delta_n} t^{(k)}(z \mid v)t(w \mid z)dzdwq(u,v)dudv
\end{aligned}
$$

By Lebesgue density theorem, this multiple integral converges to $q(x,y)t^{(k+1)}(y \mid y)$. Since I_{n2} is a sum of a finite number of terms, we conclude that I_{n2} converges to zero and this completes the proof of the lemma. □

Lemma 2.2.7 Assume the conditions of Lemma 2.2.6. Let $\{k_n\text{:n=1,2,...}\}$ be a sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. Define

$$
Z_n = \frac{1}{2\delta_n\sqrt{\lambda k_n q(x,y)}}\sum_{i=1}^{k_n}(\Psi_{ni} - \lambda\gamma(A_n \times B_n))
$$

If $n\delta_n^2 \to \infty$, then $Z_n \xrightarrow{d} N(0,1)$

Proof:

$$
\begin{aligned}
E_\Delta \Psi_{n1} &= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)I_{B_n}(X_{j+1}) \\
&= \lambda \int\int I_{A_n}(u)I_{B_n}(v)q(u,v)dudv \\
&= \lambda\gamma(A_n \times B_n)
\end{aligned}
$$

By Lemma 2.2.6, $\mathrm{Var}_\Delta(Z_n) \to 1$. So it is enough to check Lindeberg's condition: Let us write

$$
\begin{aligned}
Z_n &= \frac{1}{2\delta_n\sqrt{\lambda k_n q(x,y)}}\sum_{i=1}^{k_n}(\Psi_{ni} - \lambda\gamma(A_n \times B_n)) \\
&= \sum_{i=1}^{k_n} W_{ni} \qquad \text{(say)}
\end{aligned}
$$

Then

$$
\begin{aligned}
L_n(\epsilon) &= \sum_{i=1}^{k_n} E_\Delta(W_{ni}^2 : |W_{ni}| > \epsilon) \\
&= k_n E_\Delta(W_{n1}^2 : |W_{n1}| > \epsilon)
\end{aligned}
$$

Now, observe that

$$
\begin{aligned}
\alpha_n &\stackrel{\text{def}}{=} E_\Delta\left[(\Psi_{n1} - \lambda\gamma(A_n \times B_n))^2 : |\Psi_{n1} - \lambda\gamma(A_n \times B_n)| > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right] \\
&\leq 2\left[E_\Delta\{\Psi_{n1}^2 + (\lambda\gamma(A_n \times B_n))^2\} : |\Psi_{n1} - \lambda\gamma(A_n \times B_n)| > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right] \\
&\leq 2E_\Delta\left[\Psi_{n1}^2 : |\Psi_{n1} - \lambda\gamma(A_n \times B_n)| > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right] + 2(\lambda\gamma(A_n \times B_n))^2 \\
&= 2E_\Delta\left[\Psi_{n1}^2 : \Psi_{n1} > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)} + \lambda\gamma(A_n \times B_n)\right] \\
&\quad +2E_\Delta\left[\Psi_{n1}^2 : \Psi_{n1} < -2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)} + \lambda\gamma(A_n \times B_n)\right] + 2(\lambda\gamma(A_n \times B_n))^2 \\
&= \alpha_{n1} + \alpha_{n2} + \alpha_{n3} \qquad\qquad \text{(say)}
\end{aligned}
$$

By the Lebesgue density theorem, $\alpha_{n3} = O(\delta_n^4)$. The first term, α_{n1}, is bounded by

$$
\begin{aligned}
& E_\Delta\left[\Psi_{n1}^2 : \Psi_{n1} > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right] \\
=\; & E_\Delta\left[\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_{j+1}) : \Psi_{n1} > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right] \\
& +2E_\Delta \\
& \left[\sum_{j=1}^{T_\Delta^{(1)}-1} I_{A_n}(X_j) I_{B_n}(X_{j+1}) \sum_{k=j+1}^{T_\Delta^{(1)}-1} I_{A_n}(X_k) I_{B_n}(X_{k+1}) : \Psi_{n1} > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right] . \\
=\; & \alpha_{n11} + \alpha_{n12} \qquad\qquad \text{(say)}
\end{aligned}
$$

Now $\frac{1}{4\delta_n^2}\alpha_{n12}$ is bounded by I_n of Lemma 2.2.6 which converges to zero. Next, α_{n11} is bounded above by

$$
E_\Delta\left[\Psi_{n1} : \Psi_{n1} > 2\epsilon\sqrt{\lambda k_n \delta_n^2 q(x,y)}\right]
$$

$$
\begin{aligned}
&\leq \left[E_\Delta(\Psi_{n1}^2)E_\Delta(I^2(\Psi_{n1} > 2\epsilon\sqrt{\lambda k_n\delta_n^2 q(x,y)}))\right]^{\frac{1}{2}} \qquad \text{(by Schwartz inequality)} \\
&= \left[E_\Delta(\Psi_{n1}^2)P_\Delta(\Psi_{n1} > 2\epsilon\sqrt{\lambda k_n\delta_n^2 q(x,y)}\right]^{\frac{1}{2}} \\
&\leq \left[E_\Delta(\Psi_{n1}^2)\frac{E_\Delta(\Psi_{n1}^2)}{2\epsilon\lambda k_n\delta_n^2 q(x,y)}\right]^{\frac{1}{2}} \qquad \text{(by Chebishev inequality)} \\
&= O(\frac{\delta_n^2}{\sqrt{2\epsilon\lambda k_n\delta_n^2 q(x,y)}}) \\
&= o(\delta_n^2) \qquad \text{(since } k_n\delta_n^2 \to \infty)
\end{aligned}
$$

Thus, $\alpha_{n1} = o(\delta_n^2)$ and similarly $\alpha_{n2} = o(\delta_n^2)$ implying $\alpha_n = o(\delta_n^2)$. Thus

$$
\begin{aligned}
L_n(\epsilon) &= \frac{k_n}{4\delta_n^2\lambda^2 k_n q(x,y)}\alpha_n \\
&= o(1) \qquad \square
\end{aligned}
$$

Lemma 2.2.8 Assume the conditions of Lemma 2.2.6. Let $n\delta_n^2 \to \infty$. Let $\{K_n\text{:n=1,2,...}\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p. 1, $0 < \alpha < \infty$. Then

$$
\frac{1}{2\delta_n\sqrt{\lambda K_n q(x,y)}}\sum_{i=1}^{K_n}(\Psi_{ni} - \lambda\gamma(A_n \times B_n)) \xrightarrow{d} N(0,1)
$$

The proof is similar to the proof of Lemma 2.1.3.

Theorem 2.2.9 Under conditions of Lemma 2.2.8,

$$
\sqrt{n\delta_n^2}\left[q_n(x,y) - \frac{\gamma(A_n \times B_n)}{4\delta_n^2}\right] \xrightarrow{d} N(0,\frac{1}{4}q(x,y))
$$

Proof:

$$
\begin{aligned}
&\sqrt{n\delta_n^2}\left[q_n(x,y) - \frac{\gamma(A_n \times B_n)}{4\delta_n^2}\right] \\
&= \sqrt{n\delta_n^2}\left[\frac{1}{4n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni} - \frac{\gamma(A_n \times B_n)}{4n\delta_n^2}\right]
\end{aligned}
$$

$$+\sqrt{n\delta_n^2}\left[\frac{1}{4n\delta_n^2}\left(\sum_{j=0}^{T_\Delta^{(1)}-1} I_{A_n}(X_j)I_{B_n}(X_{j+1}) + \sum_{j=T_\Delta^{(k_n)}}^{n} I_{A_n}(X_j)I_{B_n}(X_{j+1})\right)\right]$$

It is enough to show that the first term converges in distribution to the desired limit. By Lemma 2.2.8,

$$\frac{1}{2\sqrt{\lambda K_n\delta_n^2}}\sum_{i=1}^{K_n}(\Psi_{ni}-\lambda\gamma(A_n\times B_n)) \xrightarrow{d} N(0,q(x,y))$$

Since $\frac{K_n}{n}\to\lambda^{-1}$, this implies that

$$\frac{1}{4\sqrt{n\delta_n^2}}\sum_{i=1}^{K_n}(\Psi_{ni}-\lambda\gamma(A_n\times B_n)) \quad \xrightarrow{d} \quad N(0,\frac{q(x,y)}{4})$$

The last expression is equal to

$$\begin{aligned}
\frac{\sqrt{n\delta_n^2}}{4n\delta_n^2}\sum_{i=1}^{K_n}(\Psi_{ni}-\lambda\gamma(A_n\times B_n)) &= \sqrt{n\delta_n^2}\left[\frac{1}{4n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni}-\frac{K_n}{n}\frac{\lambda\gamma(A_n\times B_n)}{4\delta_n^2}\right]\\
&= \sqrt{n\delta_n^2}\left[\frac{1}{4n\delta_n^2}\sum_{i=1}^{K_n}\Psi_{ni}-\frac{\gamma(A_n\times B_n)}{4\delta_n^2}\right]\\
&\quad +\frac{\gamma(A_n\times B_n)}{4\delta_n^2}\delta_n\sqrt{n}\left(\frac{1}{\lambda}-\frac{K_n}{n}\right)\lambda
\end{aligned}$$

By Lemma 2.1.8, the second term converges to zero in probability and this completes the proof. $\square$

Now, observe that

$$\begin{aligned}
\sqrt{n\delta_n^2}(q_n(x,y)-q(x,y)) &= \sqrt{n\delta_n^2}\left[q_n(x,y)-\frac{\gamma(A_n\times B_n)}{4\delta_n^2}\right]\\
&\quad +\sqrt{n\delta_n^2}\left[\frac{\gamma(A_n\times B_n)}{4\delta_n^2}-q(x,y)\right]\\
\frac{\gamma(A_n\times B_n)}{4\delta_n^2}-q(x,y) &= \int_{A_n}\frac{1}{2\delta_n}\int_{B_n}\frac{1}{2\delta_n}(q(u,v)-q(x,y))dudv
\end{aligned}$$

If $q(x,y)$ admits Taylor expansion up to two terms of the form

$$q(u,v)-q(x,y) \quad = \quad c_1(x,y)(u-x)+c_2(x,y)(v-y)+c_3(x,y)(u-x)^2$$

$$+c_4(x,y)(v-y)^2 + c_5(x,y)(u-x)(v-y)$$
$$+o(|x-u|^2) + o(|y-v|^2),$$

then

$$\begin{aligned}\frac{\gamma(A_n \times B_n)}{4\delta_n^2} - q(x,y) &= \int_{A_n} \frac{1}{2\delta_n} \int_{B_n} \frac{1}{2\delta_n} \left[c_3(x,y)(u-x)^2 + c_4(x,y)(v-y)^2\right] dudv \\ &= O(\delta_n^2)\end{aligned}$$

So,

$$\sqrt{n\delta_n^2}\left[\frac{\gamma(A_n \times B_n)}{4\delta_n^2} - q(x,y)\right] = O(\sqrt{n\delta_n^2\delta_n^2})$$

and we have proved

Corollary 2.2.10 Let $n\delta_n^p \to 0$ for some $p : 1 < p \leq 6$, and let q admit a Taylor expansion of order 2, then

$$\sqrt{n\delta_n^2}(q_n(x,y) - q(x,y)) \xrightarrow{d} N(0, \frac{1}{4}q(x,y)).$$

We finish this section proving asymptotic normality of t_n. By Theorem 2.1.4, $p_n(x) \to f(x)$ in probability. By Corollary 2.2.10, $\sqrt{n\delta_n^2}(q_n(x,y) - q(x,y)) \xrightarrow{d} N(0, \frac{1}{4}q(x,y))$. Then if $f(x) > 0$ we have that

$$U_n \stackrel{\text{def}}{=} \sqrt{n\delta_n^2}\left[\frac{q_n(x,y)}{p_n(x)} - \frac{q(x,y)}{p_n(x)}\right] \xrightarrow{d} N(0, \frac{q(x,y)}{4f^2(x)})$$

Also, note that

$$U_n = \sqrt{n\delta_n^2}\left[\frac{q_n(x,y)}{p_n(x)} - \frac{q(x,y)}{f(x)}\right] + \sqrt{n\delta_n^2}\left[\frac{q(x,y)}{f(x)} - \frac{q(x,y)}{p_n(x)}\right]$$

The second term can be written as $\frac{\sqrt{\delta_n}q(x,y)}{p_n(x)f(x)}\sqrt{n\delta_n}(p_n(x) - f(x))$. So, by using Theorem 2.1.9 we have proved

Theorem 2.2.11 Under conditions of Theorem 2.1.4, Corollary 2.2.10 and Theorem 2.1.9,

$$\sqrt{n\delta_n^2}(t_n(y \mid x) - t(y \mid x)) \xrightarrow{d} N\left(0, \frac{q(x,y)}{4f^2(x)}\right).$$

2.3 The General Case

In this section we consider the general case in which there is not a recurrence point Δ but the chain is recurrent as in Definition 1.3.5. Assume first $n_0 = 1$. In this situation, there exist a set $A \in \Sigma$, a probability measure ϕ on A, and a real number $\epsilon > 0$ such that:

(i) $P_x(\tau_A < \infty) = P_x(X_n \in A \text{ for some } n \geq 1) = 1 \; \forall x \in S$,

(ii) $P_x(X_1 \in E) = P(x, E) \geq \epsilon\phi(E) \; \forall x \in A$ and $\forall E \subset A$.

Following Athreya and Ney (1978), if $X_k \in A$ for $k \geq 0$, randomize the next transition as follows:

(a) With probability p $(0 < p < \epsilon)$ choose X_{k+1} over A according to ϕ,

(b) With probability $(1-p)$ choose X_{k+1} over the entire real line according to a transition function $Q(x, .)$, chosen so that the overall transition probability function for the chain remains unchanged. This is possible by (ii) above and it is achieved by taking Q so that

$$P(x, E) \;=\; p\phi(E \cap A) + (1-p)Q(x, E), \qquad x \in A, \; E \in \Sigma$$

Since A is visited infinitely often, and each time there is a probability $p > 0$ that at the next time A is entered according to ϕ, this event will ultimately occur at some time $N < \infty$ w.p. 1. This N will be the regeneration time. That is, N will play the role of T_Δ in our previous discussion. In this situation we still have independent cycles, so all the results established in sections 2.1 and 2.2 hold and the proofs are similar.

When $n_0 > 1$, (a) and (b) are changed to:

(a') With probability p $(0 < p < \epsilon)$, choose X_{k+n_0} aver A according to ϕ,

(b') With probability $(1-p)$ choose X_{k+n_0} over the entire real line according to a transition function $Q^{(n_0)}(x,.)$ chosen so that the overall transition probability for the chain remains unchanged. In this case take $Q^{(n_0)}$ such that

$$P^{(n_0)}(x,E) \;=\; p\phi(E\cap A)+(1-p)Q^{(n_0)}(x,E), \qquad x\in A,\, E\in\Sigma$$

After that , realize the segment $\{X_{k+s} : 0 < s < n_0\}$ by choosing it according to the conditional distribution of $\{X_s : 0 < s < n_0\}$ given that the boundary values X_0 and X_{n_0} are X_k and X_{k+n_0} respectively. In this case we do not have independent cycles anymore; the cycles are 1-dependent. See Asmussen (1987).

To prove consistency, we decompose the nonnegligible part of the sum, i.e., $\sum_{j=N^{(1)}}^{N(K_n)-1} I_{A_n}(X_j)$, defining the corresponding estimator in two partial sums. In the first one we include the odd cycles and in the second one we include the even cycles. For illustration, let us prove Lemma 2.1.2:

In the present situation, $\eta_{ni} = \sum_{j=N^{(i)}}^{N^{(i+1)}-1} I_{A_n}(X_j)$ and

$$\frac{1}{2k_n\delta_n}\sum_{i=1}^{k_n}\eta_{ni} \;=\; \frac{1}{2k_n\delta_n}\sum_{i=1}^{k_n}(\eta_{ni}-E_\phi\eta_{ni})+\frac{1}{2\delta_n}E_\phi\eta_{n1}$$

By Theorem 1.3.11,

$$\begin{aligned}\frac{1}{2\delta_n}E_\phi\eta_{n1} &= \frac{1}{2\delta_n}E_\phi\sum_{j=0}^{N-1} I_{A_n}(X_j)\\ &= \lambda\int_{A_n}\frac{1}{2\delta_n}f(u)du\end{aligned}$$

and by Theorem 1.4.3, this integral converges to $\lambda f(x)$.

Now,

$$\begin{aligned}\frac{1}{2k_n\delta_n}\sum_{i=1}^{k_n}(\eta_{ni}-E_\phi\eta_{ni}) &= \frac{1}{2k_n\delta_n}\sum_{i:iodd}(\eta_{ni}-E_\phi\eta_{ni})\\ &\quad+\frac{1}{2k_n\delta_n}\sum_{i:ieven}(\eta_{ni}-E_\phi\eta_{ni})\end{aligned}$$

Each partial sum on the right side is a sum of independent random variables, so it follows from earlier arguments that both converge to zero in probability . Hence by Slutsky's Theorem, the sum on the left side converges to zero in probability.

The proof of asymptotic normality offers some difficulties. We need to verify a Lindeberg condition for a double array of 1-dependent random variables. We have not completed this work yet and it will be part of future research.

2.4 Simulation

We have seen that

$$\mathrm{Var}_\Delta(p_n(x)) = \frac{f(x)}{2n\delta_n} + o(\frac{1}{n\delta_n})$$

We also know that $p_n(x)$ is biased. The bias is given by

$$B_n^2(x) = [E_\pi p_n(x) - f(x)]^2$$

If f admits Taylor expansion,

$$B_n^2(x) = \frac{1}{36}\delta_n^4\left[f''(x)\right]^2 + o(\delta_n^4)$$

So

$$\begin{aligned}E_\pi|p_n(x)-f(x)|^2 &= \frac{f(x)}{2n\delta_n} + \frac{1}{36}\delta_n^4\left[f''(x)\right]^2 + o(\delta_n^4 + \frac{1}{n\delta_n})\\ &= \alpha(n,\delta_n,x) + o(\delta_n^4+\frac{1}{n\delta_n}) \qquad \text{(say)}\end{aligned}$$

Before doing simulation, observe that given x and n, $\alpha(n, \delta_n, x)$ is a function of δ_n. If we were interested in the optimal value of δ_n for a particular point, we would get it as the minimizing value of $\alpha(n, \delta_n, x)$. We will choose a global optimal value of δ_n and our citerion will be to minimize the integrated mean squared error.

$$\begin{aligned}\int_{-\infty}^{\infty} E_\pi|p_n(x) - f(x)|^2 dx &= \frac{1}{2n\delta_n} + \frac{1}{36}\delta_n^4 \int_{-\infty}^{\infty} |f''(x)|^2 dx \\ &\quad + o(\delta_n^4 + \frac{1}{n\delta_n}) \\ &= \alpha(n, \delta_n, f) + o(\delta_n^4 + \frac{1}{n\delta_n})\end{aligned}$$

For each n, the optimal value of δ_n is that minimizing $\alpha(n, \delta_n, f)$ and it is given by $\delta_n = \beta n^{-\frac{1}{5}}$ where

$$\beta = \left[\frac{9}{2\int_{-\infty}^{\infty} |f''(x)|^2 dx}\right]^{\frac{1}{5}}$$

This expression for β was observed by Rosenblatt (1956) for the i.i.d. case.

Data from the autorregressive process $X_{n+1} = 0.5X_n + \epsilon_{n+1}$, with ϵ_n i.i.d. and $\epsilon_1 \sim N(0,1)$,were simulated and the estimator of f was calculated for $n = 100, 200, 500, 2000$. The results are shown in Figure 2.1.

There is an inconvenience with the value given for β, i.e., it depends on the unknown density being estimated. In the i.i.d case, several efforts have been made to solve this problem, but so far as we know there does not exist an universally accepted approach. Rudemo (1982) and Bowman (1984) proposed the least-squares cross-validation method. It was shown by Hall (1983) and Stone (1984) that the optimum value δ_n^* given by the least-squares cross-validation method is a consistent estimator of the optimal bandwidth. The asymptotic normality of δ_n^* was established in Hall and Marron (1987) and Scott and Terrel (1987). From the asymptotic result, Chiu (1991) observed that the bandwidth estimate δ_n^* is subject to large sample variation and proposed a stabilized bandwidth selector method. Under commonly assumed smoothness conditions, the convergence rate of this method is faster than the convergence rate for the least-squares cross-validation method. Devroye (1994) showed that for some densities Chiu's method is not consistent.

We have not gone deeply in this problem. We are considering it for future research.

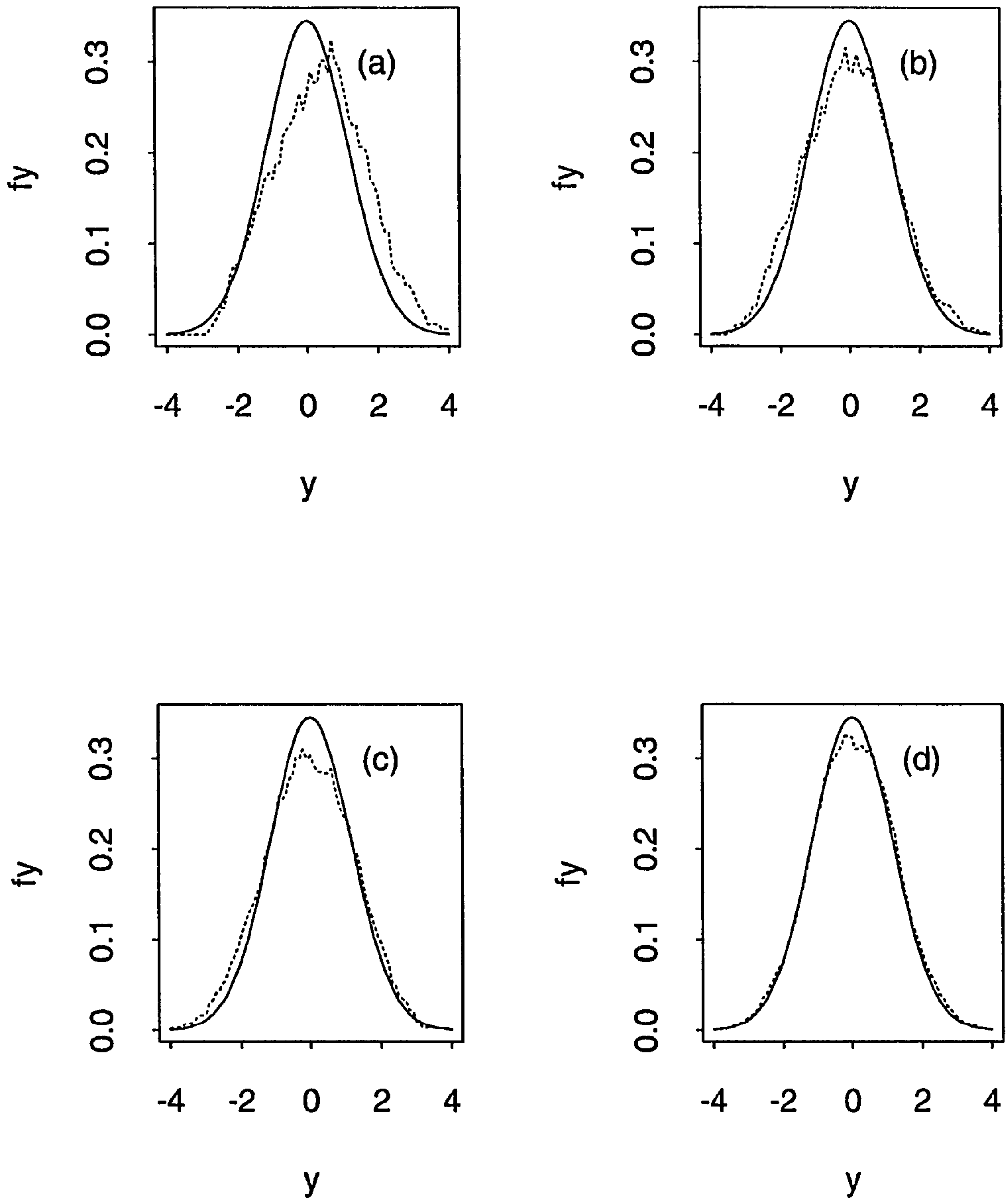

Figure 2.1: Naive K. Estimator: a:n=100, b:n=200, c:n=500, d:n=2000. Dotted line:Estimator, Continuous line: Stationary density

3. GENERAL KERNEL ESTIMATORS FOR REAL-VALUED HARRIS CHAINS

In this chapter we will extend the results of chapter 2 for kernel estimators. Now we assume that the kernel belongs to a class satisfying conditions $K1, K2, K3$ from chapter 1. Again, let T_Δ be the regeneration time. Let $\lambda = E_\Delta T_\Delta$ and assume $E_\Delta T_\Delta^2 < \infty$.

Throughout this chapter we assume that (i) $\delta_n \to 0$, (ii) The transition function $P(x, .)$ possessses a jointly continuous density $t(y \mid x)$ for all x, (iii) The invariant distribution π has a continuous density f. It is possible to relax these conditions by imposing conditions on the kernel K. However, in the interest of keeping the exposition simple we impose these strong conditions on t and f.

3.1 Kernel Estimator of the Stationary Density

Let $\pi(.)$ be the stationary probability distribution. We assume that on $R \setminus \{\Delta\}$, $\pi(.)$ is an absolutely continuous distribution with density f with respect to Lebesgue measure. When the limits of integration are not specified, the integrals are taken over the whole real line.

If the chain is observed up to time n, the kernel estimator of f at the point x is defined as

$$f_n(x) = \frac{1}{n\delta_n}\sum_{j=0}^{n} K(\frac{x-X_j}{\delta_n})$$

Our goal in this section is to study the properties of $f_n(x)$. When $K(.) = \frac{1}{2}I_{(-1,1,)}(.)$, $f_n(.)$ reduces to $p_n(.)$ of (2.1) in chapter 2.

3.1.1 Weak Consistency of f_n

The main result of this section is Theorem 3.1.4 below which establishes the weak convergence of $f_n(x)$ for $f(x)$. As in chapter 2, x will be a generic element in $R \setminus \{\Delta\}$ and all conclusions asserted are supposed to hold for almost all x (w.r.t. Lebesgue measure). The first step is the following

Lemma 3.1.1 For $i = 1, 2, ...$, define $T_\Delta^{(i)}$ as in chapter 2 and now define

$$\eta_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n})$$

Then

$$E_\Delta(\eta_{ni} - E_\Delta\eta_{ni})^2 = O(\frac{1}{\delta_n})$$

Remark. When K is the naive kernel $\frac{1}{2}I_{(-1,1)}(x)$, this η_{ni} does not reduces to η_{ni} of section (2.1) but rather to $\delta_n\eta_{ni}$ and hence $O(\frac{1}{\delta_n})$ here.

Proof: Since $\{\eta_{ni} : i = 1, 2, ...\}$ are i.i.d., it is enough to prove the result for $i = 1$.

$$\begin{aligned} E_\Delta(\eta_{n1} - E_\Delta\eta_{n1})^2 &= E_\Delta\eta_{n1}^2 - (E_\Delta\eta_{n1})^2 \\ E_\Delta\eta_{n1} &= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) \\ &= \lambda \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) f(u)du \end{aligned}$$

(the last equality by Theorem 1.3.11). By Theorem 1.4.2, the last integral converges to $\lambda f(x)$ as $\delta_n \to 0$.

$$\begin{aligned}
E_\Delta \eta_{n1}^2 &= E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) \right]^2 \\
&= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n^2} K^2(\frac{x-X_j}{\delta_n}) \\
&\quad +2E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) \frac{1}{\delta_n} K(\frac{x-X_k}{\delta_n}) \\
&= A_n + B_n \text{ (say)}
\end{aligned}$$

By Theorem 1.3.11,

$$A_n = \frac{\lambda}{\delta_n} \int \frac{1}{\delta_n} K^2(\frac{x-u}{\delta_n}) f(u) du$$

Since the last integral converges to $f(x) \int . K^2(z)dz$, we conclude that $A_n = O(\frac{1}{\delta_n})$. By Markov property and Theorem 1.3.11,

$$\begin{aligned}
B_n &= 2\lambda \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u \sum_{j=1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) f(u) du \\
&\leq \frac{M}{\delta_n} \int \frac{\lambda}{\delta_n} K(\frac{x-u}{\delta_n}) E_u(T_\Delta) f(u) du
\end{aligned}$$

(the inequality holding since the kernel is bounded by M). Now the continuity of t and f implies that $E_x(T_\Delta^{(1)})$ is continuous and by Theorem 1.4.2, the last integral converges to $\lambda E_x(T_\Delta) f(x)$. To see this, observe that $\lambda \int E_x(T_\Delta) f(x) dx = \int E_x(T_\Delta) \nu(dx)$ which is finite because $E_\Delta T_\Delta^{(2)} < \infty$. This completes the proof of the lemma. □

Lemma 3.1.2 Let $\{k_n\text{:n=1,2...}\}$ be a sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. Then, with η_{ni} as defined in Lemma 3.1.1,

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \eta_{ni} \quad \to \quad \lambda f(x) \text{ in probability}$$

provided $\delta_n \to 0$, $n\delta_n \to \infty$.

Proof:

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\eta_{ni} = \frac{1}{k_n}\sum_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni}) + E_\Delta\eta_{n1}$$

By computations in Lemma 3.1.1, $E_\Delta\eta_{n1} \to \lambda f(x)$ as $\delta_n \to 0$. So, it is enough to prove that

$$\frac{1}{k_n}\sum_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni}) \to 0 \text{ in probability:}$$

$$\begin{aligned} P_\Delta(\frac{1}{k_n}|\sum_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni})| > \epsilon) &\leq \frac{E_\Delta\left[\sum_{i=1}^{k_n}(\eta_{ni} - E_\Delta\eta_{ni})\right]^2}{k_n^2\epsilon^2} \\ &= \frac{k_n E_\Delta(\eta_{n1} - E_\Delta\eta_{n1})^2}{k_n^2\epsilon^2} \end{aligned}$$

Since $\frac{k_n}{n} \to \alpha$, the last expression converges to zero by Lemma 3.1.1 and the condition $n\delta_n \to \infty$. □

Lemma 3.1.3 Let $\{K_n$:n=1,2,...$\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w. p. 1, $0 < \alpha < \infty$. Then

$$\frac{1}{K_n}\sum_{i=1}^{K_n}\eta_{ni} \to \lambda f(x) \text{ in probability}$$

provided $\delta_n \to 0$, $n\delta_n \to \infty$.

Proof: For $\epsilon > 0$, let $A(n,\alpha,\epsilon) = \{n(\alpha - \epsilon) < K_n < n(\alpha + \epsilon)\}$. For any $\theta > 0$

$$\begin{aligned} P_\Delta(\frac{1}{K_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta) &= P_\Delta(\frac{1}{K_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta, A(n,\alpha,\epsilon)) \\ &\quad + P_\Delta(\frac{1}{K_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta, A^c(n,\alpha,\epsilon)) \\ &= \alpha_{n1} + \alpha_{n2} \qquad \text{(say)} \end{aligned}$$

Since $\frac{K_n}{n} \to \alpha$ w. p. 1, $P_\Delta(A^c(n,\alpha,\epsilon))$ and hence α_{n2} converge to zero. Now,

$$P_\Delta(\frac{1}{K_n}\sum_{i=1}^{K_n}\eta_{ni} - \lambda f(x) > \theta, A(n,\alpha,\epsilon))$$

$$\leq P_\Delta \left[\frac{1}{[n(\alpha - \epsilon)]} \sum_{i=1}^{[n(\alpha+\epsilon)]+1} \eta_{ni} - \lambda f(x) > \theta, A(n,\alpha,\epsilon) \right]$$
$$\leq P_\Delta \left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} \frac{1}{[n(\alpha+\epsilon)]+1} \sum_{i=1}^{[n(\alpha+\epsilon)]+1} \eta_{ni} - \lambda f(x) > \theta \right]$$

Since $\lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} = 1$, the last probability converges to zero by Lemma 3.1.2. Similarly it is proved that $P_\Delta(\frac{1}{K_n} \sum_{i=1}^{K_n} \eta_{ni} - \lambda f(x) < -\theta) \to 0$. □

Now, the last lemma is used to prove consistency of f_n.

Theorem 3.1.4 Let $x \neq \Delta$. If $\delta_n \to 0$, $n\delta_n \to \infty$, then

$$f_n(x) \quad \to \quad f(x) \text{ in probability}$$

Proof:

$$\begin{aligned} f_n(x) &= \frac{1}{n\delta_n} \sum_{j=0}^{n} K(\frac{x - X_j}{\delta_n}) \\ &= \frac{1}{n\delta_n} \sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n}) + \frac{1}{n} \sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} \frac{1}{\delta_n} K(\frac{x - X_j}{\delta_n}) \\ &\quad + \frac{1}{n\delta_n} \sum_{j=T_\Delta^{(K_n)}}^{n} K(\frac{x - X_j}{\delta_n}) \\ &= A_n + B_n + C_n \text{ (say)} \end{aligned}$$

where K_n is the random number of cycles, i e., the number of visits to Δ during $\{0, 1, 2, ...\}$.

Since K is bounded and $n\delta_n \to \infty$, $A_n \to 0$ w. p.1. By Lemma 2.1.8, $\frac{n - T_\Delta^{(K_n)}}{n\delta_n} \to 0$ in probability. This implies that $C_n \to 0$ in probability. So, it is enough to prove that B_n converges to $f(x)$ in probability. This follows from Lemma 3.1.3 and the fact that $\frac{K_n}{n} \to \lambda^{-1}$ w.p. 1. □

3.1.2 Asymptotic Normality of f_n

In this section we will prove asymptotic normality of f_n. The main theorems are 3.1.8 and 3.1.9. First consider the following

Lemma 3.1.5 Under the assumptions established at the begining of this chapter,

$$\delta_n E_\Delta(\eta_{ni} - E_\Delta \eta_{ni})^2 \quad \rightarrow \quad \lambda f(x) \int K^2(z)dz$$

Proof: We have already proved that $E_\Delta \eta_{n1} \rightarrow \lambda f(x)$ as $\delta_n \rightarrow 0$. Now,

$$\begin{aligned}
\delta_n E_\Delta \eta_{n1}^2 &= \delta_n E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n^2} K^2(\frac{x - X_j}{\delta_n}) \\
&\quad + 2\delta_n E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x - X_j}{\delta_n}) \frac{1}{\delta_n} K(\frac{x - X_k}{\delta_n}) \\
&= A_n + B_n \text{ (say)}
\end{aligned}$$

By Theorem 1.3.11 and Theorem 1.4.2, $A_n \rightarrow \lambda f(x) \int K^2(z)dz$ as $n \rightarrow \infty$. So, it is enough to prove that $B_n \rightarrow 0$ as $n \rightarrow \infty$. Now,

$$\begin{aligned}
B_n &= 2E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x - X_j}{\delta_n}) K(\frac{x - X_k}{\delta_n}) \right] \\
&= 2E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x - X_j}{\delta_n}) E_{X_j}(\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n})) \right] \\
&= 2\lambda \int \frac{1}{\delta_n} K(\frac{x - u}{\delta_n}) E_u \sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}) f(u)du
\end{aligned}$$

(the second equality by the Markov property and the third one by Theorem 1.3.11). Next, observe that

$$E_u \sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}) \quad = \quad E_u \left[\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}) : T_\Delta^{(1)} > j \right]$$

$$\begin{aligned}
&+E_u\left[\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n}) : T_\Delta^{(1)} \le j\right] \\
\le\ & ME_u\left[T_\Delta^{(1)} : T_\Delta^{(1)} > j\right] \\
&+\sum_{r=1}^{j} E_u\left[\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x-X_j}{\delta_n}) : T_\Delta^{(1)} = r\right]
\end{aligned}$$

So

$$\begin{aligned}
B_n \ \le\ & 2\lambda M \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u(T_\Delta^{(1)} : T_\Delta^{(1)} > j) f(u)du \\
&+2\lambda \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) \sum_{r=1}^{j} E_u\left[\sum_{k=1}^{r-1} K(\frac{x-X_k}{\delta_n}) : T_\Delta^{(1)} = r\right] f(u)du \\
=\ & B_{n1} + B_{n2} \text{ (say)}
\end{aligned}$$

It can be shown that for each j, $E_x(T_\Delta^{(1)} : T_\Delta^{(1)} > j)$ is continuous in x. Then, by Theorem 1.4.2,
$\int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u(T_\Delta^{(1)} : T_\Delta^{(1)} > j) f(u)du \to E_x(T_\Delta^{(1)} : T_\Delta^{(1)} > j) f(x)$.
This implies that

$$\begin{aligned}
&\overline{\lim}_{n\to\infty} \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u\left[\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n}) : T_\Delta^{(1)} > j\right] f(u)du \\
\le\ & ME_x(T_\Delta^{(1)} : T_\Delta^{(1)} > j) f(x).
\end{aligned}$$

Given $\epsilon > 0$, choose $j(\epsilon)$ such that $ME_x(T_\Delta^{(1)} : T_\Delta^{(1)} > j) < \epsilon$ for $j > j(\epsilon)$. Then, for $j > j(\epsilon)$,

$$\overline{\lim}_{n\to\infty} \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u\left[\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n}) : T_\Delta^{(1)} > j\right] f(u)du \ < \ \epsilon$$

Next,we will prove that $B_{n2} \to 0$ as $n \to \infty$. For each $0 \le r \le j(\epsilon)$, $0 \le k \le r$,

$$\int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u\left[K(\frac{x-X_k}{\delta_n}) : T_\Delta^{(1)} = r\right] f(u)du$$

$$
\begin{aligned}
&\le \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u K(\frac{x-X_k}{\delta_n}) f(u)du \\
&= \delta_n \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) \left[\int \frac{1}{\delta_n} K(\frac{x-z}{\delta_n}) t^{(k)}(z \mid u)dz\right] f(u)du
\end{aligned}
$$

By Theorem 1.4.2 the last integral converges to $t^{(k)}(x \mid x)f(x)$. Since $\delta_n \to 0$ and B_{n2} is a sum of a finite number of terms, we conclude that $\overline{\lim}_{n\to\infty} B_{n2} = 0$. Thus $\overline{\lim}_{n\to\infty} B_n \le \epsilon$ and ϵ being arbitrary, $\overline{\lim}_{n\to\infty} B_n = 0$ □

The Lemma 3.1.5 suggests a central limit theorem in terms of $\eta'_{ni}s$.

Lemma 3.1.6 Assume the hypothesis of Lemma 3.1.5. Let $\{k_n$:n=1,2,...$\}$ be a sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. Define

$$
Z_n = \frac{\frac{1}{\sqrt{k_n}} \sum_{i=1}^{k_n} (\eta_{ni} - E_\Delta \eta_{ni})}{\sqrt{\lambda f(x) \delta_n^{-1} \int K^2(z)dz}}
$$

If $n\delta_n \to \infty$, then $Z_n \xrightarrow{d} N(0,1)$.

Proof:

$$
\text{Var}_\Delta(Z_n) = \frac{E_\Delta(\eta_{n1} - E_\Delta \eta_{n1})^2}{\lambda f(x)\delta_n^{-1} \int K^2(z)dz} \to 1
$$

by Lemma 3.1.5. So it suffices to check Lindeberg's condition: Let us write

$$
\begin{aligned}
Z_n &= \sum_{i=1}^{k_n} \frac{\sqrt{\delta_n}(\eta_{ni} - E_\Delta \eta_{ni})}{\sqrt{\lambda f(x) k_n \int K^2(z)dz}} \\
&= \sum_{i=1}^{k_n} W_{ni} \text{ (say)}
\end{aligned}
$$

Then

$$
\begin{aligned}
L_n(\epsilon) &= \sum_{i=1}^{k_n} E_\Delta(W_{ni}^2 : |W_{ni}| > \epsilon) \\
&= k_n E_\Delta(W_{n1}^2 : |W_{n1}| > \epsilon)
\end{aligned}
$$

Let $A^2(x) = \lambda f(x) \int K^2(z)dz$.

$$\begin{aligned}
L_n(\epsilon) &= \frac{\delta_n}{A^2(x)} E_\Delta \left[(\eta_{n1} - E_\Delta \eta_{n1})^2 : |\eta_{n1} - E_\Delta \eta_{n1}| > \frac{\epsilon}{\sqrt{\delta_n}} A(x) \sqrt{k_n} \right] \\
&\leq 2 \frac{\delta_n}{A^2(x)} E_\Delta \left[\eta_{n1}^2 + (E_\Delta \eta_{n1})^2 : |\eta_{n1} - E_\Delta \eta_{n1}| > \epsilon A(x) \sqrt{k_n \delta_n^{-1}} \right] \\
&= 2 \frac{\delta_n}{A^2(x)} E_\Delta \left[\eta_{n1}^2 : |\eta_{n1} - E_\Delta \eta_{n1}| > \epsilon A(x) \sqrt{k_n \delta_n^{-1}} \right] \\
&\quad + 2 \frac{\delta_n}{A^2(x)} E_\Delta \eta_{n1}^2 P_\Delta \left[|\eta_{n1} - E_\Delta \eta_{n1}| > \epsilon A(x) \sqrt{k_n \delta_n^{-1}} \right] \\
&= \alpha_n + \beta_n \quad \text{(say)}
\end{aligned}$$

Now, we observe that by Chebyshev inequality,

$$\beta_n \leq \frac{2\delta_n}{A^2(x)} (E_\Delta \eta_{n1}^2) \frac{E_\Delta(\eta_{n1} - E_\Delta \eta_{n1})^2}{\epsilon^2 A^2(x) k_n \delta_n^{-1}}$$

So by using Lemma 3.1.1 and the fact $n\delta_n \to \infty$, we conclude that $\beta_n \to 0$. Next,

$$\frac{1}{2} A^2(x) \alpha_n = \delta_n E_\Delta(\eta_{n1}^2 : |\eta_{n1} - E_\Delta \eta_{n1}| > \epsilon A(x) \sqrt{k_n \delta_n^{-1}})$$

The last expression is, for large n, bounded above by

$$\begin{aligned}
&\delta_n E_\Delta(\eta_{n1}^2 : |\eta_{n1}| > \epsilon' A(x) \sqrt{k_n \delta_n^{-1}}) \qquad \text{(where } \epsilon' > \epsilon) \\
= \; & E_\Delta \left[\frac{1}{\delta_n} \sum_{j=0}^{T_\Delta^{(1)}-1} K^2\left(\frac{x - X_j}{\delta_n}\right) : |\eta_{n1}| > \epsilon' A(x) \sqrt{k_n \delta_n^{-1}} \right] \\
& + 2E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K\left(\frac{x - X_j}{\delta_n}\right) K\left(\frac{x - X_k}{\delta_n}\right) : |\eta_{n1}| > \epsilon' A(x) \sqrt{k_n \delta_n^{-1}} \right]
\end{aligned}$$

By the computations in the lemma above, it is enough to prove that

$$E_\Delta \left[\frac{1}{\delta_n} \sum_{j=0}^{T_\Delta^{(1)}-1} K^2\left(\frac{x - X_j}{\delta_n}\right) : |\eta_{n1} > |\epsilon' A(x) \sqrt{k_n \delta_n^{-1}} \right] \to 0$$

Since K is bounded by M,

$$
\begin{aligned}
& E_\Delta\left[\frac{1}{\delta_n}\sum_{j=0}^{T_\Delta^{(1)}-1} K^2\Big(\frac{x-X_j}{\delta_n}\Big) : |\eta_{n1}| > \epsilon' A(x)\sqrt{k_n\delta_n^{-1}}\right] \\
\leq\ & M E_\Delta\left[\frac{1}{\delta_n}\sum_{j=0}^{T_\Delta^{(1)}-1} K\Big(\frac{x-X_j}{\delta_n}\Big) : |\eta_{n1}| > \epsilon' A(x)\sqrt{k_n\delta_n^{-1}}\right] \\
\leq\ & M\left[E_\Delta\Big\{\frac{1}{\delta_n}\sum_{j=0}^{T_\Delta^{(1)}-1} K\Big(\frac{x-X_j}{\delta_n}\Big)\Big\}^2 E_\Delta\{I^2(|\eta_{n1}| > \epsilon' A(x)\sqrt{k_n\delta_n^{-1}})\}\right]^{\frac{1}{2}} \\
\leq\ & M\left[E_\Delta\Big\{\frac{1}{\delta_n}\sum_{j=0}^{T_\Delta^{(1)}-1} K\Big(\frac{x-X_j}{\delta_n}\Big)\Big\}^2 \frac{E_\Delta(\eta_{n1}^2)}{(\epsilon')^2 A^2(x) k_n\delta_n^{-1}}\right]^{\frac{1}{2}}
\end{aligned}
$$

The second inequality follows by Scwartz inequality and the third one by Chebyshev's. The last expression converges to zero, since $E_\Delta\eta_{n1}^2 = O(\frac{1}{\delta_n})$ and $k_n\delta_n \to \infty$ as $n \to \infty$. This shows that $L_n(\epsilon) \to 0$ as $n \to \infty$ and the proof of Lemma 3. 1.6 is complete. □

Lemma 3.1.7 Assume the conditions of Lemma 3.1.6. Let $n\delta_n \to \infty$. Let $\{K_n : n = 1, 2, ...\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p.1 with $0 < \alpha < \infty$. Define

$$
Z_n = \frac{\frac{1}{\sqrt{K_n}}\sum_{i=1}^{K_n}(\eta_{ni} - E_\Delta\eta_{ni})}{\sqrt{\lambda f(x)\delta_n^{-1}\int K^2(z)dz}}
$$

Then $Z_n \xrightarrow{d} N(0,1)$. The proof is similar to the proof of Lemma 3.1.3.

To finish this section we establish asymptotic normality of f_n and if $x \neq y$ we will prove that $\sqrt{n\delta_n}f_n(x)$ and $\sqrt{n\delta_n}f_n(y)$ are asymptotically independent.

Theorem 3.1.8 If $\delta_n \to 0$ and $n\delta_n \to \infty$, then

$$
\sqrt{n\delta_n}\left[f_n(x) - \frac{1}{\lambda}E_\Delta\eta_{n1}\right] \xrightarrow{d} N(0, \sigma^2(x))
$$

where $\sigma^2(x) = f(x)\int K^2(z)dz$.

Proof:

$$
\begin{aligned}
&\sqrt{n\delta_n}\left[f_n(x) - \frac{1}{\lambda}E_\Delta \eta_{n1}\right] \\
= \ & \sqrt{n\delta_n}\left[\frac{1}{n}\sum_{j=0}^{n}\frac{1}{\delta_n}K(\frac{x-X_j}{\delta_n}) - \frac{E_\Delta \eta_{n1}}{\lambda}\right] \\
= \ & \sqrt{n\delta_n}\left[\frac{1}{n\delta_n}\sum_{j=0}^{T_\Delta^{(1)}-1}K(\frac{x-X_j}{\delta_n}) + \frac{1}{n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}\frac{1}{\delta_n}K(\frac{x-X_j}{\delta_n})\right] \\
& + \sqrt{n\delta_n}\left[\frac{1}{n\delta_n}\sum_{j=T_\Delta^{(K_n)}}^{n}K(\frac{x-X_j}{\delta_n}) - \frac{E_\Delta \eta_{n1}}{\lambda}\right]
\end{aligned}
$$

Since K is bounded and $\frac{n-T^{(K_n)}}{n\delta_n}$ converges to zero in probability, it is enough to prove that

$$
Z_n = \sqrt{n\delta_n}\left[\frac{1}{n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}K(\frac{x-X_j}{\delta_n}) - \frac{E_\Delta \eta_{n1}}{\lambda}\right] \xrightarrow{d} N(0,\sigma^2(x))
$$

In terms of $\eta_{ni}'s$, Z_n can be expressed as

$$
Z_n = \sqrt{n\delta_n}(\frac{1}{n}\sum_{i=1}^{K_n}\eta_{ni} - \frac{E_\Delta \eta_{n1}}{\lambda})
$$

From Lemma 3.1.6, we know that

$$
\sqrt{K_n\delta_n}\frac{1}{K_n}\sum_{i=1}^{K_n}(\eta_{ni} - E_\Delta \eta_{ni}) \xrightarrow{d} N(0,\lambda\sigma^2(x))
$$

From this and using the fact that $\frac{K_n}{n} \to \lambda^{-1}$ w. p. 1, we conclude that

$$
W_n \stackrel{\text{def}}{=} \sqrt{\frac{\delta_n}{n}}\sum_{i=1}^{K_n}(\eta_{ni} - E_\Delta \eta_{ni}) \xrightarrow{d} N(0,\sigma^2(x))
$$

Now,

$$
W_n = \sqrt{n\delta_n}\frac{1}{n}\sum_{i=1}^{K_n}(\eta_{ni} - E_\Delta \eta_{n1})
$$

$$
\begin{aligned}
&= \sqrt{n\delta_n}\left[\frac{1}{n}\sum_{i=1}^{K_n}\eta_{ni} - \frac{K_n}{n}E_\Delta\eta_{n1}\right] \\
&= \sqrt{n\delta_n}\left[\frac{1}{n}\sum_{i=1}^{K_n}\eta_{ni} - \frac{E_\Delta\eta_{ni}}{\lambda} + (\frac{1}{\lambda} - \frac{K_n}{n})E_\Delta\eta_{n1}\right] \\
&= \sqrt{n\delta_n}\left[\frac{1}{n}\sum_{i=1}^{K_n}\eta_{ni} - \frac{E_\Delta\eta_{n1}}{\lambda}\right] + \sqrt{\delta_n}E_\Delta\eta_{n1}\sqrt{n}(\frac{1}{\lambda} - \frac{K_n}{n})
\end{aligned}
$$

By Lemma 2.1.8, the second term in the last expression converges to zero in probability and this completes the proof. □

As an immediate consequence of Theorem 3.1.8 we have

Theorem 3.1.9 Under the additional conditions $n\delta_n^p \to 0$ for some $p : 1 < p \leq 5$, K symmetric, and f twice differentiable at x,

$$\sqrt{n\delta_n}(f_n(x) - f(x)) \xrightarrow{d} N(0, \sigma^2(x))$$

Proof: Since f is twice differentiable at x,

$$
\begin{aligned}
\sqrt{n\delta_n}(f_n(x) - f(x)) &= \sqrt{n\delta_n}(f_n(x) - \frac{E_\Delta\eta_{n1}}{\lambda}) + \sqrt{n\delta_n}(\frac{E_\Delta\eta_{n1}}{\lambda} - f(x)) \\
\sqrt{n\delta_n}(\frac{E_\Delta\eta_{n1}}{\lambda} - f(x)) &= \sqrt{n\delta_n}\left[\int \frac{1}{\delta_n}K(\frac{x-u}{\delta_n})(f(u) - f(x))du\right] \\
&= \sqrt{n\delta_n}\int \frac{1}{\delta_n}K(\frac{x-u}{\delta_n})\left[f'(x)(u-x) + \frac{1}{2}f''(x)(u-x)^2\right]du \\
&\quad + o(\delta_n^2) \\
&= O(\sqrt{n\delta_n^5})
\end{aligned}
$$

So, the condition on δ_n implies that the last expression converges to zero. If K is not symmetric,we would need $n\delta_n^p \to 0$ for some $p : 1 < p \leq 3$. □

Theorem 3.1.10 Let $y \neq x$, x and $y \neq \Delta$ and consider $f_n(y) = \frac{1}{n\delta_n}\sum_{j=0}^{n} K(\frac{y-X_j}{\delta_n})$. If $\delta_n \to 0$ and $n\delta_n \to \infty$, then $\sqrt{n\delta_n}f_n(x)$ and $\sqrt{n\delta_n}f_n(y)$ are asymptotically independent.

Proof: Let

$$\begin{aligned}
\eta_{ni}(x) &= \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) \\
\tau_{ni}(y) &= \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} \frac{1}{\delta_n} K(\frac{y-X_j}{\delta_n})
\end{aligned}$$

For any l_1, l_2,

$$\begin{aligned}
& l_1\sqrt{n\delta_n}\left[f_n(x) - \frac{1}{\lambda}E_\Delta \eta_{n1}\right] + l_2\sqrt{n\delta_n}\left[f_n(y) - \frac{1}{\lambda}E_\Delta \tau_{n1}\right] \\
= \; & l_1\sqrt{n\delta_n}\left[\frac{1}{n\delta_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} K(\frac{x-X_j}{\delta_n}) - \frac{1}{\lambda}E_\Delta \eta_{n1}\right] \\
& + l_2\sqrt{n\delta_n}\left[\frac{1}{n\delta_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1} K(\frac{y-X_j}{\delta_n}) - \frac{1}{\lambda}E_\Delta \tau_{n1}\right] \\
& + l_1\frac{1}{\sqrt{n\delta_n}}\sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_j}{\delta_n}) + l_2\frac{1}{\sqrt{n\delta_n}}\sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{y-X_j}{\delta_n}) \\
& + l_1\frac{1}{\sqrt{n\delta_n}}\sum_{j=T_\Delta^{(K_n)}}^{n} K(\frac{x-X_j}{\delta_n}) + l_2\frac{1}{\sqrt{n\delta_n}}\sum_{j=T_\Delta^{(K_n)}}^{n} K(\frac{y-X_j}{\delta_n}
\end{aligned}$$

and proceeding as in the proof of Theorem 3.1.8, we can show that the above converges to a normal distribution with mean zero and variance given by

$$l_1^2 f(x)\int K^2(z)dz + l_2^2 f(y)\int K^2(z)dz + 2l_1 l_2\sigma_{12},$$

where $\sigma_{12} = lim_{n\to\infty}\delta_n \mathrm{Cov}_\Delta(\eta_{n1}, \tau_{n1})$.

Now we will show that $\sigma_{12} = 0$:

$$\begin{aligned}
\mathrm{Cov}_\Delta\left[\eta_{n1}, \tau_{n1}\right] &= E_\Delta\left[\eta_{n1}\tau_{n1}\right] - E_\Delta \eta_{n1} E_\Delta \tau_{n1} \\
E_\Delta\left[\eta_{n1}\tau_{n1}\right] &= E_\Delta\left[\sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) \sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{y-X_j}{\delta_n})\right]
\end{aligned}$$

$$
\begin{aligned}
&= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) \frac{1}{\delta_n} K(\frac{y-X_j}{\delta_n}) \\
&\quad +2E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{x-X_j}{\delta_n}) E_{X_j} \sum_{k=1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{y-X_k}{\delta_n}) \right] \\
&= A_n + B_n \qquad \text{(say)}
\end{aligned}
$$

By Theorem 1.3.11, we have that

$$
A_n = \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du
$$

Consider $x < y$ and let $\epsilon = \frac{y-x}{2}$,

$$
\begin{aligned}
A_n &= \frac{1}{\epsilon} \int_{-\infty}^{x+\epsilon} \frac{\epsilon}{\delta_n} K(\frac{x-u}{\delta_n}) \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du \\
&\quad + \frac{1}{\epsilon} \int_{x+\epsilon}^{\infty} \frac{\epsilon}{\delta_n} K(\frac{x-u}{\delta_n}) \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{1}{\epsilon} \int_{-\infty}^{x+\epsilon} \frac{y-u}{\delta_n} K(\frac{x-u}{\delta_n}) \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du \\
&\quad + \frac{1}{\epsilon} \int_{x+\epsilon}^{\infty} \frac{u-x}{\delta_n} K(\frac{x-u}{\delta_n}) \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du
\end{aligned}
$$

Given $\epsilon_1 > 0$, there exists n_0 such that on $(-\infty, x+\epsilon)$, $\frac{y-u}{\delta_n} K(\frac{y-u}{\delta_n}) < \epsilon_1$ and on $(x+\epsilon, \infty)$, $\frac{u-x}{\delta_n} K(\frac{x-u}{\delta_n}) < \epsilon_1$ for every $n \ge n_0$. Then for such n,

$$
A_n < \frac{\epsilon_1}{\epsilon} \left[\int_{-\infty}^{x+\epsilon} \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) f(u) du + \int_{x+\epsilon}^{\infty} \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du \right]
$$

$$
< \frac{\epsilon_1}{\epsilon} \left[\int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) f(u) du + \int \frac{1}{\delta_n} K(\frac{y-u}{\delta_n}) f(u) du \right]
$$

Then, $\overline{\lim}_{n\to\infty} A_n \le \frac{\epsilon_1}{\epsilon}(f(x) + f(y))$ by Theorem 1.4.2. Now, let $\epsilon_1 \to 0$. Again by Theorem 1.3.11, we have that

$$
B_n = 2\lambda \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n}) E_u \sum_{k=1}^{T_\Delta^{(1)}-1} \frac{1}{\delta_n} K(\frac{y-X_k}{\delta_n}) f(u) du
$$

$$= \frac{2\lambda}{\delta_n}\int K(\frac{x-u}{\delta_n})E_u\sum_{k=1}^{T_\Delta^{(1)}-1}\frac{1}{\delta_n}K(\frac{y-X_k}{\delta_n})f(u)du$$

As in the proof of Lemma 3.1.5, the last integral converges to zero as $n\to\infty$. Then

$$\delta_n \mathrm{Cov}_\Delta\left[\eta_{n1},\tau_{n1}\right] = \delta_n(A_n + B_n - E_\Delta\eta_{n1}E_\Delta\tau_{n1})$$

Since $E_\Delta\eta_{n1}\to\lambda f(x)$ and $E_\Delta\tau_{n1}\to\lambda f(y)$ the above goes to zero as $n\to\infty$ and this completes the proof of the Theorem. □

3.2 Kernel Estimator of the Transition Density

In this section we will prove consistency and asymptotic normality of q_n and t_n. We will assume that x is a continuity point of f and $f(x)>0$. In what follows, all assertions ares supposed to hold for almost all (x,y) with respect to Lebesgue measure in R^2 except when specific smoothness assumptions of f and q are made at particular points. By convenience in this section, $K(u,v)\stackrel{\text{def}}{=}K(u)K(v)$. So

$$q_n(x,y) = \frac{1}{n\delta_n^2}\sum_{j=0}^{n-1}K(\frac{x-X_j}{\delta_n},\frac{y-X_{j+1}}{\delta_n})$$

3.2.1 Weak Consistency of q_n and t_n

Lemma 3.2.1 With $T_\Delta^{(i)}: i=1,2,...;$ as defined in section 1, define now

$$\Psi_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1}K(\frac{x-X_j}{\delta_n},\frac{y-X_{j+1}}{\delta_n})$$

Then

$$E_\Delta(\Psi_{ni}-E_\Delta\Psi_{ni})^2 = O(\delta_n^2)$$

Proof: It is enough to prove the result for $i = 1$

$$\begin{aligned}
E_\Delta(\Psi_{n1} - E_\Delta\Psi_{n1})^2 &= E_\Delta\Psi_{n1}^2 - (E_\Delta\Psi_{n1})^2 \\
E_\Delta\Psi_{n1} &= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}) \\
&= \delta_n^2 \lambda \int\int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) q(u,v) du dv \\
&\approx \delta_n^2 \lambda q(x,y) (\int K(u) du)^2
\end{aligned}$$

(the equality by Theorem 1.3.11 and the approximation by Theorem 1.4.2). So $(E_\Delta\Psi_{n1})^2 = O(\delta_n^4)$.

$$\begin{aligned}
E_\Delta\Psi_{n1}^2 &= E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n})\right]^2 \\
&= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} K^2(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}) \\
&\quad + 2E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}) K(\frac{x - X_k}{\delta_n}, \frac{y - X_{k+1}}{\delta_n})
\end{aligned}$$

By Theorems 1.3.11 and 1.4.2 the first term is approximately equal to

$$\delta_n^2 \lambda q(x,y) (\int K^2(u) du)^2 = O(\delta_n^2)$$

By Markov property, the second term is equal to

$$\begin{aligned}
& E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}) E_{X_{j+1}} \sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}, \frac{y - X_{k+1}}{\delta_n})\right] \\
&= \lambda \int\int K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) E_v \left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}, \frac{y - X_{k+1}}{\delta_n})\right] q(u,v) du dv \\
&\leq \lambda M^2 \int\int K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) E_v(T_\Delta^{(1)}) q(u,v) du dv \\
&= \lambda M^2 \delta_n^2 \int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) E_v(T_\Delta^{(1)}) q(u,v) du dv
\end{aligned}$$

(the first equality by Theorem 1.3.11 and the inequality because K is bounded by M). To finish the proof it is enough to prove that the last integral converges to some finite number. Observe that

$$\begin{aligned}\int\int E_v(T_\Delta^{(1)}-1)q(u,v)dudv &= \int\int E_v(T_\Delta^{(1)}-1)f(u)t(u,v)dudv \\ &= \int E_v(T_\Delta^{(1)}-1)\left[\int f(u)t(u,v)du\right]dv \\ &= \int E_v(T_\Delta^{(1)}-1)f(v)dv \\ &< \infty\end{aligned}$$

By using Theorem 1.4.2, this implies that

$$\int \frac{1}{\delta_n^2}K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n})E_v(T_\Delta^{(1)}-1)q(u,v)dudv \quad\rightarrow\quad q(x,y)E_y(T_\Delta^{(1)}-1)<\infty. \quad \Box$$

Lemma 3.2.2 Let $\{k_n\text{:n=1,2,...}\}$ be a sequence of integer numbers such that $\frac{k_n}{n}\rightarrow\alpha$ with $0<\alpha<\infty$. Then

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\frac{1}{\delta_n^2}\Psi_{ni} \quad\rightarrow\quad \lambda q(x,y) \text{ in probability}$$

provided $\delta_n\rightarrow 0$ and $n\delta_n^2\rightarrow\infty$.

Proof:

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\frac{1}{\delta_n^2}\Psi_{ni} \quad=\quad \frac{1}{k_n}\sum_{i=1}^{k_n}\frac{1}{\delta_n^2}(\Psi_{ni}-E_\Delta\Psi_{ni})+\frac{1}{\delta_n^2}E_\Delta\Psi_{n1}$$

From the proof of Lemma 3.2.1, $E_\Delta\left[\frac{1}{\delta_n^2}\Psi_{n1}\right]\rightarrow\lambda q(x,y)$. So, it is enough to prove

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\frac{1}{\delta_n^2}(\Psi_{ni}-E_\Delta\Psi_{n1}) \quad\rightarrow\quad 0 \text{ in probability}$$

$$\begin{aligned}P_\Delta\left[\frac{1}{k_n\delta_n^2}\left|\sum_{i=1}^{k_n}(\Psi_{ni}-E_\Delta\Psi_{ni})\right|>\epsilon\right] &\leq \frac{E_\Delta\left[\sum_{i=1}^{k_n}(\Psi_{ni}-E_\Delta\Psi_{n1})\right]^2}{k_n\epsilon^2\delta_n^2} \\ &= \frac{1}{k_n^2\epsilon^2\delta_n^4}\sum_{i=1}^{k_n}E_\Delta(\Psi_{ni}-E_\Delta\Psi_{ni})^2 \\ &= \frac{1}{k_n\delta_n^4}E_\Delta(\Psi_{n1}-E_\Delta\Psi_{n1})^2\end{aligned}$$

The first equality because $\Psi_{ni} : i = 1, 2, ..., k_n$ are independent and the second one because they are identically distributed. By hypothesis and by Lemma 3.2.1 the last expression is $O(\frac{1}{n\delta_n^2})$ and since $n\delta_n^2 \to \infty$, the convergence in probability is proved. □

Lemma 3.2.3 Let $\{K_n$:n=1,2,...$\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p. 1 where $0 < \alpha < \infty$. Then

$$\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} \quad \to \quad \lambda q(x,y) \text{ in probability}$$

provided $\delta_n \to 0$ and $n\delta_n^2 \to \infty$.

Proof: Let $\epsilon > 0$, $\theta > 0$ and define $A(n,\alpha,\epsilon) = \{n(\alpha-\epsilon) < K_n < n(\alpha+\epsilon)\}$.

$$\begin{aligned} P_\Delta\Big(\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta\Big) &= P_\Delta\Big(\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta, A(n,\alpha,\epsilon)\Big) \\ &\quad + P_\Delta\Big(\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta, A^c(n,\alpha,\epsilon)\Big) \end{aligned}$$

Since $\frac{K_n}{n} \to \alpha$ w. p. 1, the second term converges to zero. Now,

$$\begin{aligned} & P_\Delta\Big(\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta, A(n,\alpha,\epsilon)\Big) \\ \leq\ & P_\Delta\left[\frac{1}{[n(\alpha-\epsilon)]}\sum_{i=1}^{[n(\alpha+\epsilon)]}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta, A(n,\alpha,\epsilon)\right] \\ =\ & P_\Delta\left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]}\frac{1}{[n(\alpha+\epsilon)]+1}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta, A(n,\alpha,\epsilon)\right] \\ \leq\ & P_\Delta\left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]}\frac{1}{[n(\alpha+\epsilon)]+1}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) > \theta\right] \end{aligned}$$

Since $\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} = 1$, the last probability converges to zero by Lemma 3.2.2. A similar argument proves that

$$P_\Delta\Big(\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \lambda q(x,y) < -\theta\Big) \quad \to \quad 0.$$

and this completes the proof $\square$

To finish the first part of this section we will use these lemmas to prove consistency of $q_n(x,y)$ and $t_n(y \mid x)$.

Theorem 3.2.4 If $\delta_n \to 0$ and $n\delta_n^2 \to \infty$, then

$$q_n(x,y) \quad \to \quad q(x,y) \text{ in probability}$$

Proof:

$$\begin{aligned} q_n(x,y) &= \frac{1}{n}\sum_{j=0}^{n}\frac{1}{\delta_n^2}K(\frac{x-X_j}{\delta_n},\frac{y-X_{j+1}}{\delta_n}) \\ &= \frac{1}{n\delta_n^2}\sum_{j=0}^{T_\Delta^{(1)}-1}K(\frac{x-X_j}{\delta_n},\frac{y-X_{j+1}}{\delta_n}) \\ &\quad +\frac{1}{n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni}+\frac{1}{n\delta_n^2}\sum_{j=T_\Delta^{(K_n)}}^{n}K(\frac{x-X_j}{\delta_n},\frac{y-X_{j+1}}{\delta_n}) \end{aligned}$$

Where K_n is the random number of cycles. Since K is bounded, the first term converges to zero with probability one. The third term converges to zero in probability because $\frac{n-T_\Delta^{(k_n)}}{n\delta_n^2} \to 0$ in probability. So it is enough to prove that the second term converges to $q(x,y)$ in probability.

$$\frac{1}{n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} \quad = \quad \frac{K_n}{n}\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni}$$

Since $\frac{K_n}{n} \to \lambda^{-1}$ with probability one, the result follows from Lemma 3.2.3. $\square$

Theorem 3.2.5 If $f(x) > 0$, $\delta_n \to 0$ and $n\delta_n^2 \to \infty$, then

$$t_n(y \mid x) \quad \to \quad t(y \mid x) \text{ in probability}$$

Proof: It follows by using Slutsky's theorem, Theorem 3.1.4 and Theorem 3.2.4.

3.2.2 Asymptotic Normality of q_n and t_n

In the rest of this chapter we will discuss asymptotic normality of q_n and t_n. First consider the following

Lemma 3.2.6 If $t(y \mid x)$ is continuous in x and y, then

$$\frac{1}{\delta_n} E_\Delta(\Psi_{n1} - E_\Delta \Psi_{n1})^2 \quad \rightarrow \quad \lambda q(x.y)(\int K^2(z)dz)^2$$

provided $\delta_n \rightarrow 0$ and $n\delta_n^2 \dashrightarrow \infty$.

Proof: From the proof of Lemma 3.2.2, $\frac{1}{\delta_n^2}(E_\Delta \Psi_{n1})^2 \rightarrow 0$

$$\begin{aligned} \frac{1}{\delta_n^2} E_\Delta \Psi_{n1}^2 &= \frac{1}{\delta_n^2} E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} K^2(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}) \\ &+ \frac{2}{\delta_n^2} E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} K(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}) K(\frac{x - X_k}{\delta_n}, \frac{y - X_{k+1}}{\delta_n}) \end{aligned}$$

The first term is equal to $\lambda \int\int \frac{1}{\delta_n^2} K^2(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) q(u,v)dudv$ and this integral converges to $\lambda q(x,y)(\int K^2(z)dz)^2$. So, the Lemma will be proved if we prove that the second term converges to zero. By Markov property, this sum is equal to
$\frac{2}{\delta_n^2} E_\Delta \left[\sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_j}{\delta_n}, \frac{y-X_{j+1}}{\delta_n}) E_{X_{j+1}} \sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n}, \frac{y-X_{k+1}}{\delta_n})\right]$
and by Theorem 1.3.11, equal to
$$2\lambda \int\int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) E_v \left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}) K(\frac{y - X_{k+1}}{\delta_n})\right] q(u,v)dudv$$
Now,

$$\begin{aligned} & E_v \sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}, \frac{y - X_{k+1}}{\delta_n}) \\ = \; & E_v \left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x - X_k}{\delta_n}, \frac{y - X_{k+1}}{\delta_n}) : T_\Delta^{(1)} > j\right] \end{aligned}$$

$$+E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n},\frac{y-X_{k+1}}{\delta_n}) : T_\Delta^{(1)} \le j\right]$$

$$\le \; M^2 E_v(T_\Delta^{(1)} : T_\Delta^{(1)} > j)$$

$$+\sum_{r=1}^{j} E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_j}{\delta_n},\frac{y-X_{k+1}}{\delta_n}) : T_\Delta^{(1)} = r\right]$$

Then, introducing the notation $q_{uv} \equiv q(u,v)dudv$, the last integral is bounded above by

$$M^2 \int\int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n}) E_v(T_\Delta^{(1)} : T_\Delta^{(1)} > j) q_{uv}$$

$$+\int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n}) \sum_{r=1}^{j} E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n},\frac{y-X_{k+1}}{\delta_n}) : T = r\right] q_{uv}$$

$= I_{n1} + I_{n2}$ (say)

Note that I_{n1} converges to $M^2 E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > j) q(x,y)$, then

$$\overline{\lim}_{n\to\infty} \int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n}) E_v\left[\sum_{k=1}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n},\frac{y-X_{k+1}}{\delta_n}) : T_\Delta^{(1)} > j\right] q_{uv}$$

$$\le M^2 E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > j) q(x,y)$$

Given $\epsilon > 0$, choose $j(\epsilon)$ such that $M^2 E_y(T_\Delta^{(1)} : T_\Delta^{(1)} > j) q(x,y) < \epsilon$ for $j > j(\epsilon)$. Then

$$\overline{\lim}_{n\to\infty} \int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n}) E_v\left[\sum_{k=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_k}{\delta_n},\frac{y-X_{k+1}}{\delta_n}) : T_\Delta^{(1)} > j\right] q_{uv} \le \epsilon \; .$$

Now, for each $0 \le r \le j(\epsilon)$, $0 \le k \le r$,

$$\int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n}) E_v\left[K(\frac{x-X_k}{\delta_n},\frac{y-X_{k+1}}{\delta_n}), T_\Delta^{(1)} = r\right] q_{uv}$$

$$\le \; \int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n}) E_v\left[K(\frac{x-X_k}{\delta_n},\frac{y-X_{k+1}}{\delta_n})\right] q_{uv}$$

$$= \delta_n^2 \int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) \left[\int \frac{1}{\delta_n^2} K(\frac{x-z}{\delta_n}, \frac{y-w}{\delta_n}) t^{(k)}(z \mid v) t(w \mid z) dz dw \right] q_{uv}$$

Since the last multiple integral converges to $q(x,y)t^{(k)}(y \mid y)$ and I_{n2} is a sum of a finite number of terms, we conclude that $\overline{\lim}_{n\to\infty} I_{n2} \le \epsilon$ and since $\epsilon > 0$ is arbitrary this completes the proof of the Lemma. □

Lemma 3.2.7 Let $\{k_n : n = 1, 2, ...\}$ be a sequence of integer numbers such that $\frac{k_n}{n} \to \alpha$ with $0 < \alpha < \infty$. Define

$$Z_n = \frac{1}{\sqrt{k_n \lambda q(x,y)} \int K^2(z) dz} \sum_{i=1}^{k_n} \frac{1}{\delta_n} (\Psi_{ni} - E_\Delta \Psi_{ni})$$

Then, under hypothesis of Lemma 3.2.6, $Z_n \xrightarrow{d} N(0,1)$

Proof: By Lemma 3.2.6, $\mathrm{Var}_\Delta(Z_n) \to 1$ as $n \to \infty$. So it is enough to check Lindeberg's condition: Let $A^2(x,y) = \lambda q(x,y)(\int K^2(z)dz)^2$. With this notation,

$$\begin{aligned} Z_n &= \sum_{i=1}^{k_n} \frac{1}{\sqrt{k_n} A(x,y)} \frac{1}{\delta_n} (\Psi - E_\Delta \Psi_{ni}) \\ &= \sum_{i=1}^{k_n} W_{ni} \qquad \text{(say)} \end{aligned}$$

Then

$$\begin{aligned} L_n(\epsilon) &= \sum_{i=1}^{k_n} E_\Delta(W_{ni}^2 : |W_{ni}| > \epsilon) \\ &= k_n E_\Delta(W_{n1}^2 : |W_{n1}| > \epsilon) \\ &= \frac{1}{\delta_n^2 A(x,y)} E_\Delta \left[(\Psi_{n1} - E_\Delta \Psi_{n1})^2 : |\Psi - E_\Delta \Psi_{n1}| > \epsilon \sqrt{k_n \delta_n^2} A(x,y) \right] \\ &\le \frac{2}{A(x,y)} \left[E_\Delta (\frac{\Psi_{n1}}{\delta_n})^2 + (\frac{E_\Delta \Psi_{n1}}{\delta_n})^2 : \frac{|\Psi_{n1} - E_\Delta \Psi_{n1}|}{\delta_n} > \epsilon \sqrt{k_n} A(x,y) \right] \end{aligned}$$

Since $\left[\frac{E_\Delta \Psi_{n1}}{\delta_n}\right]^2 \to 0$, it is enough to prove that $E_\Delta \left[\frac{\Psi_{n1}^2}{\delta_n^2} : \frac{|\Psi_{n1} - E_\Delta \Psi_{n1}|}{\delta_n} > \epsilon \sqrt{k_n} A(x,y) \right]$

converges to zero.

$$\begin{aligned}
& E_\Delta\left[\frac{\Psi_{n1}^2}{\delta_n^2} : \frac{|\Psi_{n1} - E_\Delta \Psi_{n1}|}{\delta_n} > \epsilon\sqrt{k_n}A(x,y)\right] \\
= \; & E_\Delta\left[\frac{\Psi_{n1}^2}{\delta_n^2} : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y) + \frac{E_\Delta \Psi_{n1}}{\delta_n}\right] \\
& + E_\Delta\left[\frac{\Psi_{n1}^2}{\delta_n^2} : \frac{\Psi_{n1}}{\delta_n} < -\epsilon\sqrt{k_n}A(x,y) + \frac{E_\Delta \Psi_{n1}}{\delta_n}\right]
\end{aligned}$$

Now observe that

$$\begin{aligned}
& E_\Delta\left[\frac{\Psi_{n1}^2}{\delta_n^2} : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y) + \frac{E_\Delta \Psi_{n1}}{\delta_n}\right] \\
\leq \; & E_\Delta\left[\frac{\Psi_{n1}^2}{\delta_n^2} : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y)\right] \\
= \; & E_\Delta\left[\frac{1}{\delta_n^2}\sum_{j=0}^{T_\Delta^{(1)}-1} K^2\left(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}\right) : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y)\right] + o(1)
\end{aligned}$$

(From the proof of Lemma 3.2.6)

To finish the proof, observe that

$$\begin{aligned}
& E_\Delta\left[\frac{1}{\delta_n^2}\sum_{j=0}^{T_\Delta^{(1)}-1} K^2\left(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}\right) : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y)\right] \\
\leq \; & M^2 E_\Delta\left[\frac{1}{\delta_n^2}\sum_{j=0}^{T_\Delta^{(1)}-1} K\left(\frac{x - X_j}{\delta_n}, \frac{y - X_{j+1}}{\delta_n}\right) : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y)\right] \\
= \; & M^2 E_\Delta\left[\frac{\Psi_{n1}}{\delta_n^2} : \frac{\Psi_{n1}}{\delta_n} > \epsilon\sqrt{k_n}A(x,y)\right] \\
\leq \; & M^2\left[\frac{E_\Delta \Psi_{n1}^2}{\delta_n^4}\frac{E_\Delta \Psi_{n1}^2}{k_n\delta_n^2\epsilon^2 A^2(x,y)}\right]^{\frac{1}{2}} \\
= \; & M^2\frac{E_\Delta \Psi_{n1}^2}{\delta_n^2}\left[\frac{1}{k_n\delta_n^2\epsilon^2 A^2(x,y)}\right]^{\frac{1}{2}}
\end{aligned}$$

Since $k_n\delta_n^2 \to \infty$, the result follows from Lemma 3.2.6. $\square$

Lemma 3.2.8 Lemma 3.2.7 holds with $\{K_n : n = 1, 2, ...\}$ a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p.1 with $0 < \alpha < \infty$. The proof is similar to the proof of Lemma 3.2.3.

Theorem 3.2.9 If $t(y \mid x)$ is continuous in x and y and $n\delta_n^2 \to \infty$, then

$$\sqrt{n\delta_n^2}(q_n(x,y) - \frac{E_\Delta \Psi_{n1}}{\lambda \delta_n^2}) \xrightarrow{d} N(0, \sigma^2(x,y))$$

where $\sigma^2(x,y) = q(x,y)(\int K^2(z)dz)^2$.

Proof:

$$\begin{aligned}
&\sqrt{n\delta_n^2}(q_n(x,y) - \frac{E_\Delta \Psi_{n1}}{\lambda \delta_n^2}) \\
= \;& \sqrt{n\delta_n^2}\left[\frac{1}{n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \frac{E_\Delta \Psi_{n1}}{\lambda\delta_n^2}\right] \\
&+\sqrt{n\delta_n^2}\left[\frac{1}{n\delta_n^2}\sum_{j=0}^{T_\Delta^{(1)}-1} K(\frac{x-X_j}{\delta_n}, \frac{y-X_{j+1}}{\delta_n}) + \frac{1}{n\delta_n^2}\sum_{j=T_\Delta^{(K_n)}}^{n} K(\frac{x-X_j}{\delta_n}, \frac{y-X_{j+1}}{\delta_n})\right]
\end{aligned}$$

Since K is bounded and $\frac{n-T_\Delta^{K_n)}}{n\delta_n^2} \to 0$ in probability, it is enough to prove that

$$\sqrt{n\delta_n^2}\left[\frac{1}{n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \frac{E_\Delta \Psi_{n1}}{\lambda\delta_n^2}\right] \xrightarrow{d} N(0, \sigma^2(x,y))$$

We know from Lemma 3.2.8 that

$$W_n \stackrel{\text{def}}{=} \sqrt{\lambda K_n \delta_n^2}\frac{1}{\lambda K_n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}(\Psi_{ni} - E_\Delta \Psi_{ni}) \xrightarrow{d} N(0, \sigma^2(x,y))$$

But

$$\begin{aligned}
W_n &= \sqrt{\delta_n^2}\frac{1}{\sqrt{\lambda K_n}}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}(\Psi_{ni} - E_\Delta \Psi_{ni}) \\
&= \sqrt{\frac{\delta_n^2}{n}}\sqrt{\frac{n}{\lambda K_n}}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}(\Psi_{ni} - E_\Delta \Psi_{n1})
\end{aligned}$$

Since $\frac{K_n}{n} \to \lambda^{-1}$ w.p. 1, we conclude that

$$T_n \stackrel{\text{def}}{=} \sqrt{\frac{\delta_n^2}{n}}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}(\Psi_{ni} - E_\Delta \Psi_{n1}) \xrightarrow{d} N(0, \sigma^2(x,y))$$

Now, observe that

$$\begin{aligned} T_n &= \sqrt{n\delta_n^2}\left[\frac{1}{n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \frac{K_n}{n}\frac{E_\Delta\Psi_{n1}}{\delta_n^2}\right] \\ &= \sqrt{n\delta_n^2}\left[\frac{1}{n}\sum_{i=1}^{K_n}\frac{1}{\delta_n^2}\Psi_{ni} - \frac{E_\Delta\Psi_{n1}}{\lambda\delta_n^2}\right] + \delta_n\frac{E_\Delta\Psi_{n1}}{\delta_n^2}\sqrt{n}(\frac{1}{\lambda} - \frac{K_n}{n}) \end{aligned}$$

Since $\frac{E_\Delta\Psi_{n1}}{\delta_n^2} \to \lambda q(x,y)(\int K(u)du)^2$ (from the proof of Lemma 3.2.1), Lemma 2.1.8 completes the proof. □

Now observe that

$$\begin{aligned} \sqrt{n\delta_n^2}(q_n(x,y) - q(x,y)) &= \sqrt{n\delta_n^2}\left[q_n(x,y) - \frac{E_\Delta\Psi_{n1}}{\lambda\delta_n^2}\right] \\ &\quad + \sqrt{n\delta_n^2}\left[\frac{E_\Delta\Psi_{n1}}{\lambda\delta_n^2} - q(x,y)\right] \end{aligned}$$

If $q(x,y)$ admits Taylor expansion,

$$\begin{aligned} q(u,v) - q(x,y) &= c_1(x,y)(u-x) + c_2(x,y)(v-y) + \frac{1}{2}c_3(x,y)^2 \\ &\quad + \frac{1}{2}c_4(x,y)^2 + c_5(u-x)(v-y) + o(\delta_n^2) \end{aligned}$$

Then if K is symmetric,

$$\begin{aligned} &\frac{E_\Delta\Psi_{n1}}{\lambda\delta_n^2} - q(x,y) \\ &= \int \frac{1}{\delta_n^2}K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n})(q(u,v) - q(x,y))dudv \\ &= \frac{1}{2}\int \frac{1}{\delta_n^2}K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n})\left[c_3(x,y)(u-x)^2 + c_4(x,y)(v-y)^2\right]dudv \\ &\quad + o(\delta_n^2) \end{aligned}$$

Now,

$$\int \frac{1}{\delta_n^2}K(\frac{x-u}{\delta_n},\frac{y-v}{\delta_n})(u-x)^2 dudv$$

$$
\begin{aligned}
&= \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n})(u-x)^2 \left[\int \frac{1}{\delta_n} K(\frac{y-v}{\delta_n}) dv\right] du \\
&= \int \frac{1}{\delta_n} K(\frac{x-u}{\delta_n})(u-x)^2 du \\
&= \delta_n \int (\frac{u-x}{\delta_n})^2 K(u-x) du \\
&= \delta_n^2 \int z^2 K(z) dz
\end{aligned}
$$

Similarly, it is proved that $\int \frac{1}{\delta_n^2} K(\frac{x-u}{\delta_n}, \frac{y-v}{\delta_n}) dudv = O(\delta_n^2)$

So

$$
\begin{aligned}
\sqrt{n\delta_n^2}\left[\frac{E_\Delta \Psi_{n1}}{\lambda \delta_n^2} - q(x,y)\right] &= O(\sqrt{n\delta_n^2}\delta_n^2) \\
&= O(\sqrt{n\delta_n^6})
\end{aligned}
$$

Then we have proved the following

Theorem 3.2.10 Let q be of class C^3 in a neighborhood of (x, y). If K is symmetric and $n\delta_n^p \to 0$ for some $2 < p \leq 6$, then

$$
\sqrt{n\delta_n^2}(q_n(x,y) - q(x,y)) \xrightarrow{d} N(0, \sigma^2(x,y))
$$

where $\sigma^2(x,y) = q(x,y)(\int K^2(z)dz)^2$.

To finish this chapter, we prove asymptotic normality of t_n:

Theorem 3.2.11 If $f(x) > 0$ then under conditions of Theroem 3.2.10 and Theorem 3.1.4,

$$
\sqrt{n\delta_n^2}(t_n(y \mid x) - t(y \mid x)) \xrightarrow{d} N(0, \tau(x,y))
$$

where $\tau(x,y) = \frac{\sigma^2(x,y)}{f^2(x)} = \frac{q(x,y)\left[\int K^2(z)dz\right]^2}{f^2(x)}$

Proof: By Theorem 3.2.10 and Theorem 3.1.4,

$$
\sqrt{n\delta_n^2}\left[\frac{q_n(x,y)}{f_n(x)} - \frac{q(x,y)}{f_n(x)}\right] \xrightarrow{d} N\left(0, \frac{\sigma^2(x,y)}{f^2(x)}\right)
$$

Now, observe that

$$\begin{aligned}
\sqrt{n\delta_n^2}\left[\frac{q_n(x,y)}{f_n(x)}-\frac{q(x,y)}{f_n(x)}\right] &= \sqrt{n\delta_n^2}\left[\frac{q_n(x,y)}{f_n(x)}-\frac{q(x,y)}{f(x)}\right] \\
&\quad +\sqrt{n\delta_n^2}\left[\frac{q(x,y)}{f(x)}-\frac{q(x,y}{f_n(x)}\right] \\
&= \sqrt{n\delta_n^2}\left[t_n(y\mid x)-t(y\mid x)\right] \\
&\quad +\frac{\sqrt{\delta_n}q(x,y)}{f_n(x)f(x)}\sqrt{n\delta_n}(f_n(x)-f(x))
\end{aligned}$$

By Theorem 3.1.9, the second term converges to zero in probability and this completes the proof. □

3.3 Simulation

For the general kernel estimator , the variance of $f_n(x)$ is given by

$$\mathrm{Var}_\Delta f_n(x)=\frac{f(x)\int K^2(z)dz}{n\delta_n}+o(\frac{1}{n\delta_n})$$

If f admits Taylor expansion, the bias is given by

$$\begin{aligned}
B_n^2(x) &= \left[E_\pi f_n(x)-f(x)\right]^2 \\
&= \left[-\delta_n f'(x)\int zK(z)dz+\frac{1}{2}\delta_n^2 f''(x)\int z^2K(z)dz\right]^2+o(\delta_n^4)
\end{aligned}$$

If K is symmetric,

$$B_n^2(x)=\frac{1}{4}\delta_n^4\left[f''(x)\int z^2K(z)dz\right]^2+o(\delta_n^4)$$

So

$$\begin{aligned}
\int E_\pi|f_n(x)-f(x)|^2dx &= \frac{1}{n\delta_n}\int K^2()z)dz+\frac{1}{4}\delta_n^4\left[\int z^2K(z)dz\right]^2\int|f''(x)|^2dx \\
&\quad +o(\frac{1}{n\delta_n}+\delta_n^4) \\
&= \alpha(K,n,\delta_n,f)+o(\frac{1}{n\delta_n}+\delta_n^4)
\end{aligned}$$

Given n and the kernel K, $\alpha(K, n, \delta_n, f)$ is a function in δ_n (and f). The optimal value of δ_n is that minimizing it and it is given by $\delta_n = \beta n^{-\frac{1}{5}}$ where now,

$$\beta = \left[\frac{\int K^2(z)dz}{\left[\int z^2 K(z)dz\right]^2 \int |f''(x)|^2 dx} \right]^{\frac{1}{5}}$$

Using the data simulated in chapter 2, it was estimated f by using the standard normal kernel. The results are shown in Figure 3.1 .

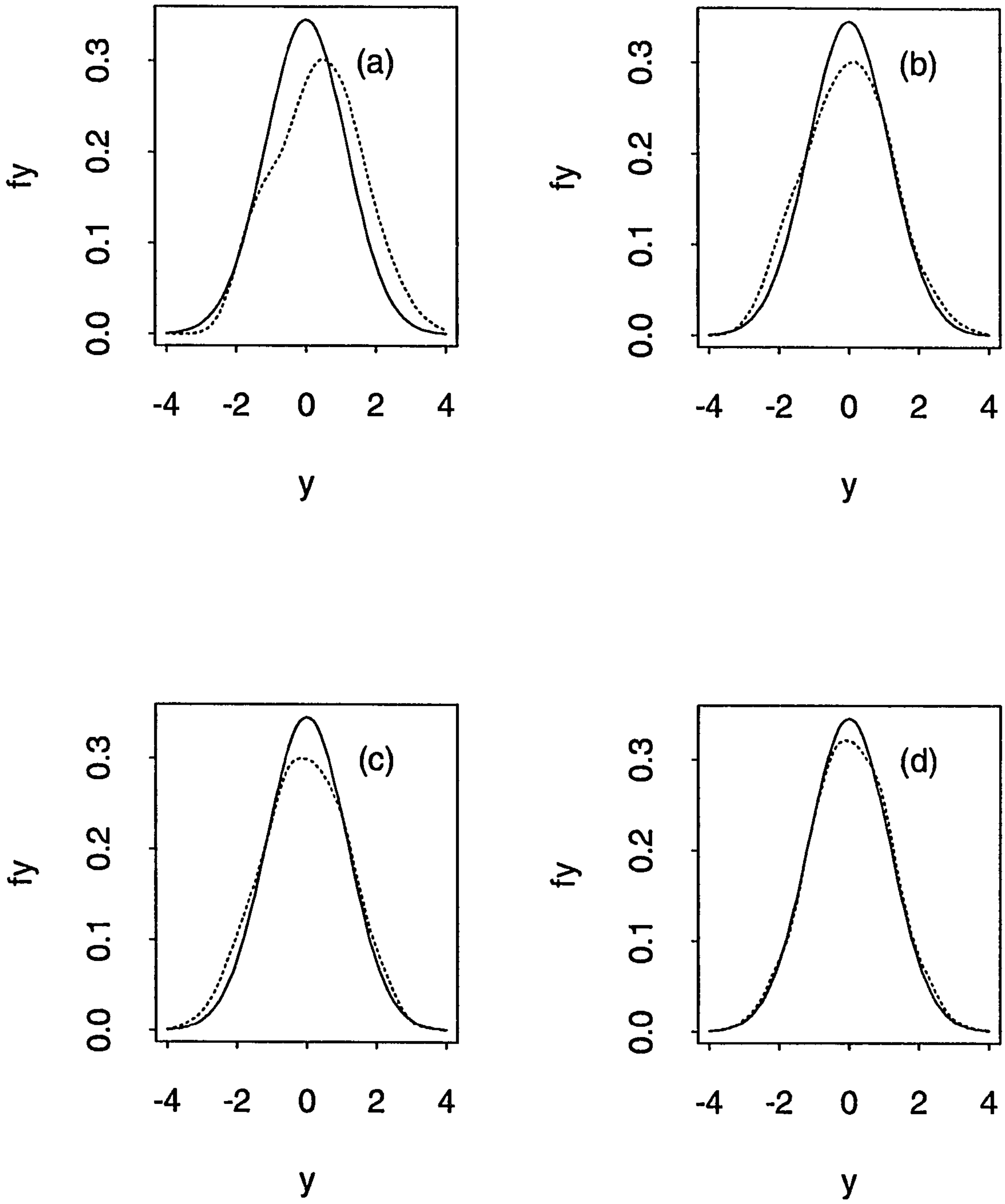

Figure 3.1: Normal Estimator: a:n=100, b:n=200, c:n=500, d:n=2000. Dotted line:Estimator, Continuous line:Stationary Density

4. KERNEL ESTIMATORS FOR SEMI-MARKOV PROCESSES

4.1 Introduction

Let $\{X_n : n = 0, 1, ...\}$ be a Markov chain with state space (S, Σ). Assume that there exists a $\Delta \in S$ such that $P_x(T_\Delta < \infty) = 1$ and $E_\Delta T_\Delta < \infty$. Recall that if $\{X_n : n = 0, 1, ...\}$ is a Harris chain then for some $n_0 \geq 1$, $\{X_{nn_0} : n = 0, 1, ...\}$ is a Markov chain for which a distinguished point Δ can be constructed by enlarging the state space. Thus the results of this chapter apply to Harris chains.

Let $\{G(x, .) : x \in S\}$ be a family of distribution functions such that $G(x, 0) = 0$ for all $x \in S$. Given $\{X_n : n = 0, 1, ...\}$, let $\{T_n : n = 0, 1, ...\}$ be a sequence of independent random variables such that

$$P(T_n \leq t \mid \{X_n : n = 0, 1, ...\}) = G(X_n, t) \qquad \text{for all } t \geq 0.$$

These $\{T_n : n = 0, 1, ...\}$ are called sojourn times. Let $S_0 = 0$, $S_n = \sum_{i=0}^{n-1} T_i$ for $n \geq 1$ and let

$$Y_t = X_n \qquad S_n \leq t < S_{n+1},\ n = 0, 1, 2, ...$$

Since $P_x(X_n = \Delta \text{ for some } n \geq 1) = 1$, $\{X_n : n = 0, 1, ...\}$ hits Δ infinitely often, say at $N_1, N_2, ...$ and so by the strong law of large numbers, $\sum_i T_{N_i} = \infty$ w.p. 1 and hence $S_n \to \infty$ w.p. 1 and so $X(t)$ is well defined for all t. The process $\{Y_t : t \geq 0\}$

has state space S and is not Markov unless $G(x,.)$ is exponential for all x. However, $\{X_{T_n} = X_n : n = 0,1,...\}$ is still a Markov chain. Thus $\{X_t : t \geq 0\}$ sampled at $t = T_n$, $n = 0,1,2,...$ is Markov but not for all $t \geq 0$. For this reason it is called a **semi-Markov process.**

In this chapter, we assume $S = R$ and $\Sigma = B(R)$.

Estimators for the stationary density f and the transition density t of the Markov chain $\{X_n : n = 0,1,...\}$ are the same as defined in previous chapters.

The main goals in this chapter will be to propose estimator for $G(x,t)$ and prove its properties.

With this in mind, as before, we observe the process up to time n. Besides the information $\{X_0, X_1, ..., X_n\}$, now we have $\{T_0, T_1, ..., T_{n-1}\}$ where for $i = 0,1,...,n-1$, T_i is the sojourn time in X_i.

Fix $x \in R \setminus \{\Delta\}$ and take $A_n = (x - \delta_n, x + \delta_n)$. During $\{0,1,...,n\}$ let $N_1, N_2, ..., N_{L_n}$ be the times of visits by the process to A_n; i.e., for $i = 1,2,...,L_n$, $X_{N_i} \in A_n$; and let T_{N_i} be the corresponding sojourn times. Given $X_0, X_1, ..., X_n$; $T_{N_i} : i = 1,2,...,L_n$ are independent, but not identically distributed.

Define

$$G_n(x,t) = \frac{1}{L_n} \sum_{i=1}^{L_n} I(T_{N_i} \leq t)$$

We propose $G_n(x,t)$ as an estimator of $G(x,t)$. We will prove consistency of $G_n(x,t)$ in section 2 and asymptotic normality in section 3. The method of proof of these results is the same one used in previous chapters.

4.2 Consistency of $G_n(x,t)$

The main result of this section is Theorem 4.2.3 establishing weak consistency of $G_n(x,t)$.

Lemma 4.2.1 Let $\{k_n : n = 1, 2, ...\}$ be a sequence of integers such that $\frac{k_n}{n} \to \alpha$, $0 < \alpha < \infty$. For $i = 1, 2, ..., k_n$, define

$$\xi_{ni} = \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} \frac{1}{2\delta_n} |G(X_j,t) - G(x,t)| I(X_j \in A_n)$$

If $\lim_{\delta_n \to 0} \int_{|x-y|<\delta_n} \frac{1}{2\delta_n} |G(y,t) - G(x,t)| f(y) dy = 0$, then

$$\overline{\xi}_n \equiv \frac{1}{k_n} \sum_{i=1}^{k_n} \xi_{ni} \to 0 \text{ in probability}$$

Proof: Since $\{\xi_{ni} : i = 1, 2, ...\}$ are i.i.d and $E_\Delta \xi_{n1} \to 0$ by hypothesis, $E_\Delta \overline{\xi}_n \to 0$ and since $\overline{\xi}_n \geq 0$, the result follows. □

Lemma 4.2.2 Let $\{K_n : n = 1, 2, ...\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \to \alpha$ w.p. 1 with $0 < \alpha < \infty$. Under the hypothesis of Lemma 4.2.1,

$$\frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} \to 0 \qquad \text{in probability}$$

Proof: Let $\epsilon > 0$, $\theta > 0$ and let $A(n, \alpha, \epsilon) = \{n(\alpha - \epsilon) < K_n < n(\alpha + \epsilon)\}$

$$\begin{aligned} P_\Delta \left[\left| \frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} \right| > \theta \right] &= P_\Delta \left[\frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} > \theta \right] + P_\Delta \left[\frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} < -\theta \right] \\ &= a_{n1} + a_{n2} \qquad \text{(say)} \end{aligned}$$

$$\alpha_{n1} = P_\Delta \left[\frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} > \theta, A(n,\alpha,\epsilon) \right] + P_\Delta \left[\frac{1}{K_n} \sum_{i=1}^{K_n} \xi_{ni} > \theta, A^c(n,\alpha,\epsilon) \right]$$

Since $\frac{K_n}{n} \to \alpha$ w.p. 1; the second term converges to zero. Now,

$$\begin{aligned}
& P_\Delta\left[\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right] \\
\leq \; & P_\Delta\left[\frac{1}{[n(\alpha-\epsilon)]}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right] \\
= \; & P_\Delta\left[\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon]}\frac{1}{[n(\alpha+\epsilon)]+1}\sum_{i=1}^{[n(\alpha+\epsilon)]+1}\xi_{ni} > \theta, A(n,\alpha,\epsilon)\right]
\end{aligned}$$

Since $\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{[n(\alpha+\epsilon)]+1}{[n(\alpha-\epsilon)]} = 1$ as $n \to \infty$, the last probability converges to zero by Lemma 4.2.1. Similarly it is proved that $\alpha_{n2} \to 0$. □

Theorem 4.2.3 Fix $x \in R\setminus\{\Delta\}$. Let $f(x) > 0$. Let $\delta_n \to 0$. If $\lim_{\delta_n\to 0}\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}|G(y,t)-G(x,t)|f(y)dy = 0$, $\frac{1}{\delta_n}\int_{A_n}f(y)dy \to f(x)$, and $n\delta_n \to \infty$ as $n \to \infty$, then

$$G_n(x,t) \;\to\; G(x,t) \qquad \text{in probability}$$

Proof: Recall that $N_1, N_2, ..., N_{L_n}$ are times of visits to A_n and

$$\begin{aligned}
G_n(x,t) &= \frac{1}{L_n}\sum_{j=1}^{L_n} I(T_{N_j} \leq t) \\
&= \frac{1}{L_n}\sum_{j=1}^{L_n}[I(T_{N_j} \leq t) - G(X_{N_j},t)] \\
&\quad + \frac{1}{L_n}\sum_{j=1}^{L_n}[G(X_{N_j},t) - G(x,t)] + G(x,t) \\
&= \alpha_{n1} + \alpha_{n2} + G(x,t)
\end{aligned}$$

It was proved in chapter 2 that if $\frac{1}{\delta_n}\int_{A_n}f(y)dy \to f(x)$ then $p_n(x) \to f(x)$ in probability (Theorem 2.1.4). But from definition, $p_n(x) = \frac{L_n}{2n\delta_n}$. So, $L_n \to \infty$ in probability if $f(x) > 0$. Besides that, conditioned on $\{X_n : n = 0, 1, ...\}$, $(I(T_{N_j} \leq t) - G(X_{N_j}, t))$ are independent random variables with mean zero. Then α_{n1} converges to zero in probability. Next we will prove that α_{n2} converges to zero in probability:

$$
\begin{aligned}
\frac{1}{L_n}\sum_{j=1}^{L_n}[G(X_{N_j},t)-G(x,t)] &\le \frac{1}{L_n}\sum_{j=0}^{n}|G(X_j,t)-G(x,t)|I(X_j\in A_n) \\
&= \frac{1}{L_n}\sum_{j=0}^{T_\Delta^{(1)}-1}|G(X_j,t)-G(x,t)|I(X_j\in A_n) \\
&\quad +\frac{1}{L_n}\sum_{T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}|G(X_j,t)-G(x,t)|I(X_j\in A_n) \\
&\quad +\frac{1}{L_n}\sum_{T_\Delta^{(K_n)}}^{n}|G(X_j,t)-G(x,t)|I(X_j\in A_n) \\
&= \alpha_{n21}+\alpha_{n22}+\alpha_{n23} \qquad \text{(say)}
\end{aligned}
$$

Since $L_n \to \infty$ in probability, α_{n21} converges to zero in probability. Also α_{n23} converges to zero in probability because $\frac{n-T_\Delta^{(K_n)}}{L_n}$ does by thightness of $(n-T_\Delta^{(K_n)})$.

Now,

$$
\begin{aligned}
\alpha_{n22} &= \frac{K_n}{n}\frac{2n\delta_n}{L_n}\frac{1}{K_n}\sum_{i=1}^{K_n}\sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1}\frac{1}{2\delta_n}|G(X_j,t)-G(x,t)|I(X_j\in A_n) \\
&= \frac{K_n}{n}\frac{2n\delta_n}{L_n}\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}
\end{aligned}
$$

Therefore, by Lemma 2.1.8, Theorem 2.1.4, and Lemma 4.2.2, α_{n22} converges to zero in probability. □

4.3 Asymptotic Normality of $G_n(x,t)$

The main result of this section is Theorem 4.3.5.

Lemma 4.3.1 For $i = 1, 2, \ldots$; define

$$
\Psi_{ni} = \frac{1}{2\delta_n}\sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1} G(X_j,t)(1-G(X_j,t))I(X_j\in A_n)
$$

Let $\frac{1}{\delta_n}\int_{A_n} E_u T_\Delta f(u)du$ and $\frac{1}{\delta_n}\int_{A_n} f(u)du$ be bounded in n. Then

$$E_\Delta(\Psi_{n1} - E_\Delta \Psi_{n1})^2 \;=\; O(\frac{1}{\delta_n})$$

Proof:

$$\begin{aligned}
E_\Delta \Psi_{n1}^2 &= \frac{1}{4\delta_n^2} E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} G^2(X_j,t)(1-G(X_j,t))^2 I(X_j \in A_n) \\
&\quad + \frac{1}{2\delta_n^2} E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} \sum_{k=j+1}^{T_\Delta^{(1)}-1} G(X_j,t)(1-G(X_j,t))G(X_k,t) \\
&\qquad (1-G(X_k,t))I(X_j \in A_n)I(X_k \in A_n) \\
&= \alpha_n + \beta_n \qquad \text{(say)}
\end{aligned}$$

$$\begin{aligned}
\alpha_n &= \frac{\lambda}{2\delta_n}\int_{A_n} \frac{1}{2\delta_n} G^2(u,t)(1-G(u,t))^2 f(u)du \\
&\le \frac{\lambda}{64\delta_n}\int_{A_n} \frac{1}{\delta_n} f(u)du \\
&= O(\frac{1}{\delta_n})
\end{aligned}$$

Next,

$$\begin{aligned}
\beta_n &= \frac{1}{\delta_n} E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1} G(X_j,t)(1-G(X_j,t))I(X_j \in A_n) \\
&\qquad E_{X_j} \sum_{K=1}^{T_\Delta^{(1)}-1} \frac{1}{2\delta_n} G(X_k,t)(1-G(X_k,t))I(X_k \in A_n) \\
&\le \frac{\lambda}{\delta_n}\int_{A_n} \frac{1}{2\delta_n} G(u,t)(1-G(u,t))E_u(T_\Delta^{(1)}-1)f(u)du \\
&\le \frac{\lambda}{8\delta_n}\int_{A_n} \frac{1}{\delta_n} E_u T_\Delta f(u)du \\
&= O(\frac{1}{\delta_n}) \qquad \square
\end{aligned}$$

Lemma 4.3.2 Assume the hypotheses of Theorem 4.2.3. Let $\{k_n : n = 1,2,...\}$ be a sequence of integers such that $\frac{k_n}{n} \to \alpha$ with $0 < \alpha < \infty$. Let $\delta_n \to 0$ and $n\delta_n \to \infty$.

Then,

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\Psi_{ni} \rightarrow \lambda f(x)G(x,t)(1-G(x,t)) \qquad \text{in probability}$$

Proof:

$$\frac{1}{k_n}\sum_{i=1}^{k_n}\Psi_{ni} = \frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi_{ni}-E_\Delta\Psi_{ni}) + E_\Delta\Psi_{n1}$$

By Theorem 1.3.11, $E_\Delta\Psi_{n1} = \lambda\int_{A_n}\frac{1}{2\delta_n}G(y,t)(1-G(y,t))f(y)dy$ and under the hypotheses of Theorem 4.2.3, this integral converges to $\lambda f(x)G(x,t)(1-G(x,t))$. So, it is enough to prove that $\frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi_{ni}-E_\Delta\Psi_{ni}) \rightarrow 0$ in probability:

$$\begin{aligned} P_\Delta(\frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi - E_\Delta\Psi_{ni}) > \epsilon) &\leq \frac{1}{\epsilon^2 k_n}E_\Delta(\Psi_{n1}-E_\Delta\Psi_{n1})^2 \\ &= O(\frac{1}{k_n\delta_n}) \end{aligned}$$

Similarly, it is proved that $P_\Delta(\frac{1}{k_n}\sum_{i=1}^{k_n}(\Psi - E_\Delta\Psi_{ni}) < -\epsilon) = O(\frac{1}{k_n\delta_n})$. Since $k_n\delta_n \rightarrow \infty$, the proof is complete. □

Lemma 4.3.3 Let $\{K_n : n = 1,2,...\}$ be a sequence of integer random variables such that $\frac{K_n}{n} \rightarrow \alpha$ w.p. 1 with $0 < \alpha < \infty$. Under the hypotheses of Lemma 4.3.2,

$$\frac{1}{K_n}\sum_{i=1}^{K_n}\Psi_{ni} \rightarrow \lambda f(x)G(x,t)(1-G(x,t)) \qquad \text{in probability}$$

The proof is similar to the proof of Lemma 4.2.3.

Lemma 4.3.4 Under hypotheses of Theorem 4.3.2 and Lemma 4.3.1,

$$\frac{1}{L_n}\sum_{j=0}^{L_n}G(X_{N_j},t)(1-G(X_{N_j},t)) \rightarrow G(x,t)(1-G(x,t)) \qquad \text{in probability}$$

Proof:

$$\frac{1}{L_n}\sum_{j=0}^{L_n}G(X_{N_j},t)(1-G(X_{N_j},t))$$

$$
\begin{aligned}
&= \frac{1}{L_n}\sum_{j=0}^{n} G(X_j,t)(1-G(X_j,t))I(X_j \in A_n) \\
&= \frac{2n\delta_n}{L_n}\frac{K_n}{n}\frac{1}{K_n}\sum_{i=1}^{K_n}\frac{1}{2\delta_n}\sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(1+1)}-1} G(X_j,t)(1-G(X_j,t))I(X_j \in A_n) \\
&\quad + \frac{1}{L_n}\sum_{j=0}^{T_\Delta^{(1)}-1} G(X_j,t)(1-G(X_j,t))I(X_j \in A_n) \\
&\quad + \frac{1}{L_n}\sum_{j=T_\Delta^{(K_n)}}^{n} G(X_j,t)(1-G(X_j,t))I(X_j \in A_n) \\
&= \alpha_{n1} + \alpha_{n2} + \alpha_{n3} \qquad \text{(say)}
\end{aligned}
$$

By earlier arguments, α_{n2} and α_{n3} converge to zero in probability. Also, by Theorem 2.1.4, Lemma 2.1.8, and Lemma 4.3.3; α_{n1} converges to $G(x,t)(1-G(x,t))$ in probability and the proof is complete. $\square$

Now we establish asymptotic normality of $G_n(x,t)$

Theorem 4.3.5 Assume the hypotheses of Theorem 4.2.3 and Lemma 4.3.1, i. e., assume that:

(i) $\quad \frac{1}{\delta_n}\int_{A_n} f(y)dy$ is bounded in n,

(ii) $\quad \frac{1}{2\delta_n}\int_{A_n} |G(y,t)-G(x,t)|f(y)dy \to 0,$

(iii) $\quad \frac{1}{\delta_n}\int_{A_n} E_u T_\Delta f(u)du$ bounded in n,

(iv) $\quad \delta_n \to 0,\ n\delta_n \to \infty$ as $n \to \infty$.

Assume also that:

(v) $\quad$ There exist $\alpha > 0$ and for each t, a C_t such that

$|G(y,t)-G(x,t)| \le C_t|x-y|^{2+\alpha}$ for y near x,

(vi) $\quad n\delta_n^p \to 0$ for some $1 < p \le 5+2\alpha$.

Let

$$Z_n = \frac{\frac{1}{L_n}\sum_{j=1}^{L_n}(I(T_{X_{N_j}} \le t) - G(x,t))}{\sqrt{\frac{1}{L_n^2}\sum_{j=1}^{L_n} G(X_{N_j},t)(1-G(X_{N_j},t))}}$$

Then, $Z_n \xrightarrow{d} N(0,1)$.

We write

$$\begin{aligned} Z_n &= \frac{\frac{1}{L_n}\sum_{j=1}^{L_n}(I(T_{X_{N_j}} \le t) - G(X_{N_j},t))}{\sqrt{\frac{1}{L_n^2}\sum_{j=1}^{L_n} G(X_{N_j},t)(1-G(X_{N_j},t))}} \\ &\quad + \frac{\frac{1}{L_n}\sum_{j=1}^{L_n}(G(X_{N_j},t) - G(x,t))}{\sqrt{\frac{1}{L_n^2}\sum_{j=1}^{L_n} G(X_{N_j},t)(1-G(X_{N_j},t))}} \\ &= Z_{n1} + Z_{n2} \qquad \text{(say)} \end{aligned}$$

To prove Theorem 4.3.5, first we will prove Lemma 4.3.7 and Lemma 4.3.8 below. For Lemma 4.3.7 we need the following result from Chung(1974), pp 199.

Lemma 4.3.6 Let $\{\theta_{nj}, 1 \le j \le k_n, 1 \le n\}$ be a double array of complex numbers satisfying the following conditions as $n \to \infty$:

(i) $\max\{|\theta_{nj}| : 1 \le j \le k_n\} \to 0$;

(ii) $\sum_{j=1}^{k_n} |\theta_{nj}| \le M < \infty$, where M does not depend on n;

(iii) $\sum_{j=1}^{k_n} \theta_{nj} \to \theta$,where θ is a (finite)complex number.

Then we have

$$\prod_{j=1}^{k_n}(1+\theta_{nj}) \to \exp(\theta)$$

Lemma 4.3.7 Let $D = \sigma(X_0, X_1, ...)$. Let $\phi_n(\theta) = E_\Delta\left[\exp(i\theta Z_{n1}) \mid D\right]$. Then for each θ in R, $\phi_n(\theta) \to \exp(-\frac{1}{2}\theta^2)$ in probability.

Proof: Let $\delta_{nj} = I(T_{X_{N_j}} \le t)$, $p_{nj} = G(X_{N_j},t)$, and $v_n = \sqrt{\sum_{j=1}^{L_n} p_{nj}(1-p_{nj})}$.

For θ in R,

$$\begin{aligned}\phi_n(\theta) &= E_\Delta\left[\exp(\frac{i\theta}{v}\sum_{j=1}^{L_n}(\delta_{nj}-p_{nj})) \mid D\right] \\ &= \prod_{j=1}^{L_n} E_\Delta\left[\exp(\frac{i\theta}{v}(\delta_{nj}-p_{nj})) \mid D\right] \text{ (by conditional independence of } \delta_{nj})\end{aligned}$$

Let $a(\theta,n,j) = 1+\theta_{nj}$ where $\theta_{nj} = E_\Delta\left[\exp(\frac{i\theta}{v_n}(\delta_{nj}-p_{nj})) - 1 \mid D\right]$.

Since $E_\Delta\left[(\delta_{nj}-p_{nj}) \mid D\right] = 0$,

$$\begin{aligned}\theta_{nj} &= E_\Delta\left[\{\exp(\frac{i\theta}{v_n}(\delta_{nj}-p_{nj})) - 1 - \frac{i\theta}{v_n}(\delta_{nj}-p_{nj})\} \mid D\right] \\ &= E_\Delta\left[\{\exp(\frac{i\theta}{v_n}(\delta_{nj}-p_{nj})) - 1 - \frac{i\theta}{v_n}(\delta_{nj}-p_{nj}) - \frac{1}{2}(\frac{i\theta}{v_n})^2(\delta_{nj}-p_{nj})^2\} \mid D\right] \\ &\quad - \frac{\theta^2}{2v_n^2}p_{nj}(1-p_{nj})\end{aligned}$$

Since $|\exp(it)-1-it| \le \frac{t^2}{2!}$ and $|exp(it)-1-it-\frac{(it)^2}{2!}| \le \frac{|t|^3}{3!}$ for t real (See Feller (1971), pp 512), we have that

$$|\theta_{nj}| \le \frac{\theta^2 p_{nj}(1-p_{nj})}{2v_n^2}$$

Thus $\sum_{j=1}^{L_n}|\theta_{nj}| \le \frac{\theta^2}{2}$ and $max\{|\theta_{nj}| : 1 \le j \le L_n\} \le \frac{\theta^2}{8v_n^2}$.

But by Lemma 4.3.4, $\frac{v_n^2}{L_n} \to G(x,t)(1-G(x,t)$ in probability and since $L_n \to \infty$ in probability as well, $\max\{|\theta_{nj}| : 1 \le j \le L_n\} \to 0$ in probability under conditioning by D.

Also

$$\begin{aligned}&\sum_{j=1}^{L_n}\theta_{nj} + \frac{\theta^2}{2} \\ &= \sum_{j=1}^{L_n} E_\Delta\left[\{\exp(\frac{i\theta}{v_n}(\delta_{nj}-p_{nj}) - 1 - \frac{i\theta}{v_n}(\delta_{nj}-p_{nj}) - \frac{1}{2}(\frac{i\theta}{v_n})^2(\delta_{nj}-p_{nj})^2\} \mid D\right]\end{aligned}$$

Then

$$
\begin{aligned}
|\sum_{j=1}^{L_n} \theta_{nj} + \frac{\theta^2}{2}| &\le \frac{1}{3!}\sum_{j=1}^{L_n}(\frac{|\theta|}{v_n})^3 E_\Delta\left[(\delta_{nj} - p_{nj})^3 \mid D\right] \\
&\le 2\frac{|\theta|^3}{3!v_n^3}\sum_{j=1}^{L_n} E_\Delta\left[(\delta_{nj} - p_{nj})^2 \mid D\right] \qquad (\text{since } |\delta_{nj} - p_{nj}|^2 \le 2) \\
&\le \frac{1}{3}\frac{|\theta|^3}{v_n}
\end{aligned}
$$

Since $v_n \to \infty$ in probability, $|\sum_{j=1}^{L_n} \theta_{nj} + \frac{\theta^2}{2}| \to 0$ in probability under conditioning by D.

For any subsequence n' there exists a further subsequence n'' such that along that, with probability one: $\max\{|\theta_{nj}| : 1 \le j \le L_n\} \to 0$, $\sum_{j=1}^{L_n} \theta_{nj} + \frac{\theta^2}{2} \to 0$, and $\sum_{j=1}^{L_n} |\theta_{nj}| \le \frac{\theta^2}{2}$; and so by Lemma 4.3.6, $\phi_n(\theta) \to \exp(-\frac{\theta^2}{2})$ w.p. 1 along that subsequence n''. This being true for every subsequence n', the result follows. □

Lemma 4.3.8 Assume conditions (v) and (vi) in Theorem 4.3.5, i.e., (v) There exist $\alpha > 0$ and for each t, a C_t such that $|G(y,t) - G(x,t)| \le C_t|y - x|^{2+\alpha}$. (vi) $n\delta_n^p \to 0$ for some $1 < p \le 5 + 2\alpha$. Then $Z_{n2} \to 0$ in probability, where Z_{n2} defined in Theorem 4.3.5 is

$$
Z_{n2} = \frac{\frac{1}{\sqrt{L_n}}\sum_{j=1}^{L_n}(G(X_{N_j},t) - G(x,t))}{\sqrt{\frac{1}{L_n}\sum_{j=1}^{L_n} G(X_{N_j},t)(1 - G(X_{N_j},t))}}
$$

Proof: By Lemma 4.3.4, it is enough to prove that the numerator converges to zero in probability. To do this observe that

$$
\begin{aligned}
\frac{1}{\sqrt{L_n}}\sum_{j=1}^{L_n}(G(X_{N_j},t) - G(x,t)) &= \frac{1}{\sqrt{L_n}}\sum_{j=0}^{n}(G(X_j,t) - G(x,t))I(X_j \in A_n) \\
&= \frac{1}{\sqrt{L_n}}\sum_{j=0}^{T_\Delta^{(1)}-1}(G(X_j,t) - G(x,t))I(X_j \in A_n) \\
&\quad + \frac{1}{\sqrt{L_n}}\sum_{j=T_\Delta^{K_n}}^{n}(G(X_j,t) - G(x,t))I(X_j \in A_n)
\end{aligned}
$$

$$+\frac{1}{\sqrt{L_n}}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}(G(X_j,t)-G(x,t))I(X_j\in A_n)$$
$$=\quad \alpha_{n1}+\alpha_{n2}+\alpha_{n3}\qquad \text{(say)}$$

By earlier arguments, $\alpha_{n1}+\alpha_{n2}$ converges to zero in probability. Next,

$$\begin{aligned}\alpha_{n3} &= \frac{1}{\sqrt{L_n}}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}(G(X_j,t)-G(x,t))I(X_j\in A_n)\\ &= \frac{\sqrt{2n\delta_n f(x)}}{\sqrt{L_n}}\frac{\sqrt{2n\delta_n f(x)}}{f(x)}\frac{K_n}{n}\frac{1}{K_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n)\end{aligned}$$

By Theorem 2.1.4 and Lemma 2.1.8, it is enough to prove that

$$\sqrt{2n\delta_n}\frac{1}{K_n}\sum_{j=T_\Delta^{(1)}}^{T_\Delta^{(K_n)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n)\ \rightarrow\ 0\qquad \text{in probability}$$

To do that, consider first K_n as non random. To go to the case in which K_n is random we follow the procedure used earlier. Let

$$\xi_{ni}\ =\ \sum_{j=T_\Delta^{(i)}}^{T_\Delta^{(i+1)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n)$$

$$\begin{aligned}P_\Delta\left[\sqrt{2n\delta_n}\left|\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}\right|>\epsilon\right] &\le \frac{2n\delta_n E_\Delta(\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni})^2}{\epsilon^2}\\ &= \frac{2n\delta_n}{\epsilon^2}\left[\frac{1}{K_n}E_\Delta\xi_{n1}^2+\frac{K_n(K_n-1)}{2K_n^2}E_\Delta\xi_{n1}E_\Delta\xi_{n2}\right]\\ &= O(\delta_n E_\Delta\xi_{n1}^2)+O(n\delta_n(E_\Delta\xi_{n1})^2)\end{aligned}$$

$$\begin{aligned}E_\Delta\xi_{n1}^2 &= \frac{1}{4\delta_n^2}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}(G(X_j,t)-G(x,t))^2I(X_j\in A_n)\\ &\quad+\frac{1}{4\delta_n^2}E_\Delta\sum_{j=0}^{T_\Delta^{(1)}-1}\sum_{k=j+1}^{T_\Delta^{(1)}-1}(G(X_j,t)-G(x,t))(G(X_k,t)-G(x,t))\\ &\qquad\qquad I(X_j\in A_n)I(X_k\in A_n)\\ &= \alpha_{n31}+\alpha_{n32}\qquad \text{(say)}\end{aligned}$$

$$\begin{aligned}
\alpha_{n31} &= \frac{1}{2\delta_n}\lambda\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}(G(y,t)-G(x,t))^2 f(y)dy \\
&\leq \frac{\lambda C_t}{2\delta_n}\delta_n^{4+2\alpha}\int_{|x-y|<\delta_n} f(y)dy \\
&= O(\delta_n^{4+2\alpha})
\end{aligned}$$

By Theorem 1.4.2,

$$\begin{aligned}
\alpha_{n32} &= E_\Delta \sum_{j=0}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}(G(X_j,t)-G(x,t))I(X_j\in A_n)E_{X_j} \\
&\qquad\qquad \sum_{k=1}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}(G(X_k,t)-G(x,t))I(X_k\in A_n) \\
&= \lambda\int_{|x-y|<\delta_n}\frac{1}{2\delta_n} \\
&\quad \left[(G(y,t)-G(x,t))E_y\sum_{k=1}^{T_\Delta^{(1)}-1}\frac{1}{2\delta_n}(G(X_k,t)-G(x,t))I(X_k\in A_n)\right] f(y)dy \\
&\leq \frac{\lambda C_t}{2}\delta_n^{3+2\alpha}\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}E_y(T_\Delta^{(1)}-1)f(y)dy \\
&= O(\delta_n^{3+2\alpha})
\end{aligned}$$

Hence, $E_\Delta\xi_{n1}^2 = O(\delta_n^{3+2\alpha})$. Also we have that

$$\begin{aligned}
|E_\Delta\xi_{n1}| &= \lambda\int_{|x-y|<\delta_n}\frac{1}{2\delta_n}|G(y,t)-G(x,t)|f(y)dy \\
&\leq \frac{\lambda C_t}{2\delta_n}\int_{|x-y|<\delta_n}\delta_n^{2+\alpha}f(y)dy \\
&= O(\delta_n^{2+\alpha}) \qquad \text{(since } f \text{ is bounded)}
\end{aligned}$$

Therefore, $P_\Delta\left[\sqrt{2n\delta_n}|\frac{1}{K_n}\sum_{i=1}^{K_n}\xi_{ni}|>\epsilon\right] = O(n\delta_n^{5+2\alpha})$. $\square$

Proof of Theorem 4.3.5: For θ in R,

$$\begin{aligned}
E(\exp(i\theta Z_{n1})) &= E(E_\Delta(\exp(i\theta Z_{n1}\mid D)) \\
&= E(\phi_n(\theta))
\end{aligned}$$

By Lemma 4.3.7 and the bounded convergence theorem, the last expectation converges to $exp(-\frac{\theta^2}{2})$. Thus $Z_{n1} \xrightarrow{d} N(0,1)$.

By Lemma 4.3.8, $Z_{n2} \to 0$ in probability. So by Slutsky's theorem, $Z_n = Z_{n1} + Z_{n2} \xrightarrow{d} N(0,1)$. $\square$

Remark 1. Notice that Z_n is pivotal for $G(x,t)$ since the limit law of Z_n is $N(0,1)$ and is independent of all parameters. Thus, Z_n could be used to obtain confidence intervals for $G(x,t)$.

Remark 2. In the hypotheses of Theorem 4.3.5 (i), (ii), and (iii) hold for almost all x by Lebesgue density theorem (Theorem 1.4.3). $E_\Delta T_\Delta^2 < \infty$ is necessary for (iii).

4.4 Simulation

Take the autoregressive process from section 2.4.

Let $\{G(x,t) : x \in R\}$ be a family of exponential distributions with parameter $\lambda_x = 10 + |x|$, i.e.,

$$G(x,t) = 1 - \exp(-\frac{1}{10+|x|}t) \qquad t \geq 0.$$

Given $\{X_n : n = 0, 1, ...\}$, let $\{T_n : n = 0, 1, ...\}$ be a sequence of independent random variables such that

$$\begin{aligned} P(T_n \leq t \mid X_n : n = 0,1,...) &= G(X_n, t) \\ &= 1 - \exp(-\frac{1}{10+|X_n|}t) \qquad t \geq 0 \end{aligned}$$

For $n \geq 0$, and for $t \geq 0$ define S_n and Y_t respectively as in section 4.1.

Samples of size n=200 were generated from this process and by using the same value for δ_n used in section 2.4, we estimated $G(x,t)$ for $x = -1$, $x = -0.5$, $x = 0.5$, and $x = 1$. The results are shown in Figure 4.1.

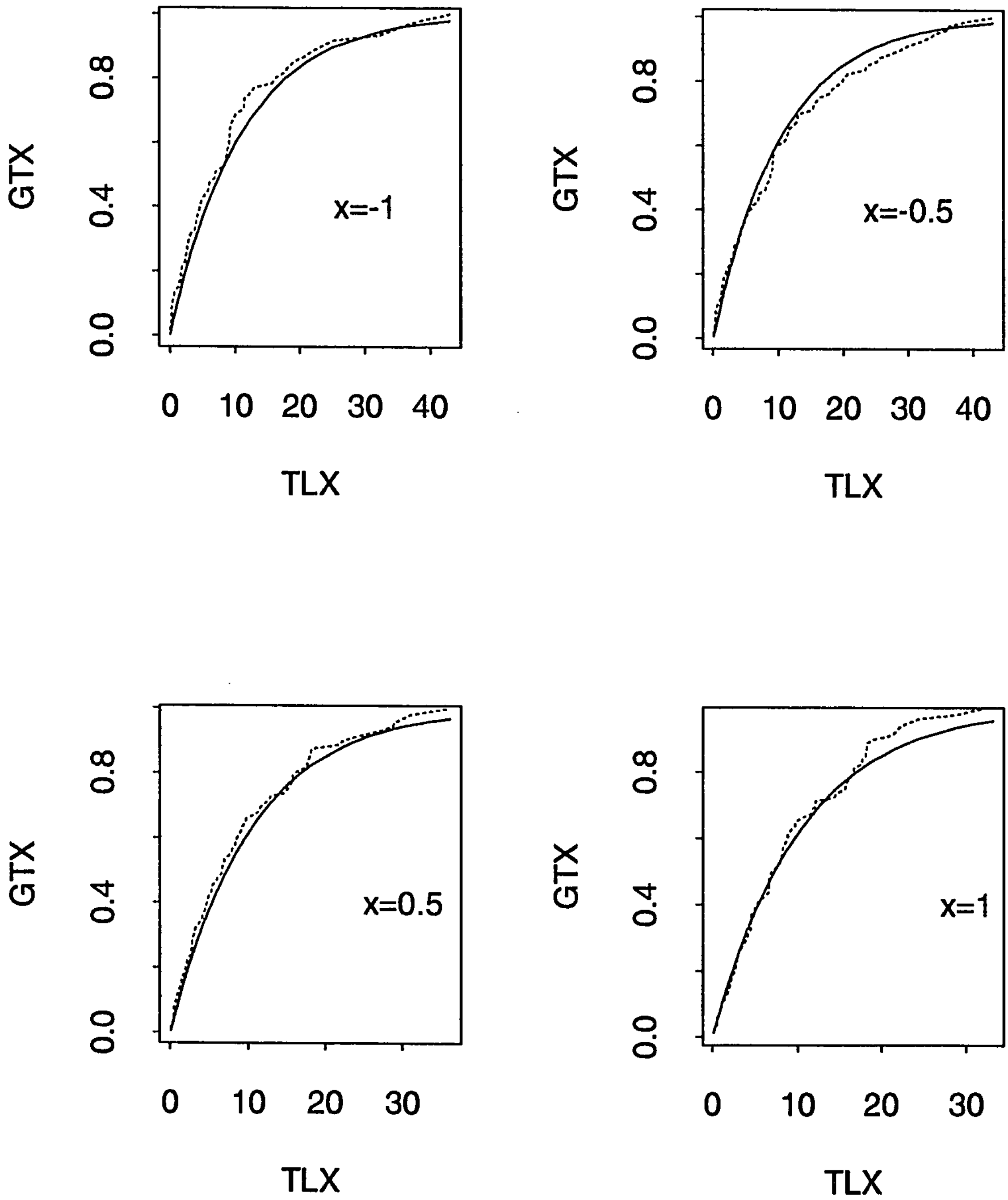

Figure 4.1: **Estimator of the sojourn time distribution (STD). Dotted line:Estimator, Continuous line:STD.**

5. CONCLUSIONS AND FURTHER RESEARCH

We have proved asymptotic results for kernel estimators of the stationary density and the transition density of a Markov chain under Harris condition. This is much weaker than Doeblin condition. Examples of Harris chains that do not satisfy Doeblin condition were given in chapter 1.

The applications of our results are immediate. In fact, we have already presented an specific application to autoregressive processes. Following with this example, we can also define an estimator for ρ, based on f_n and q_n. We suggest an estimator for ρ, given by

$$\rho_n = \frac{\int xyq_n(x,y)dxdy}{\int x^2 f_n(x)dx} .$$

In a near future we will study the properties of this estimator.

As it was said in chapter 1, our results can also be applied to the problem of storage of water in dams or reservoirs. There exists a Markov chain associated with this problem and that chain satisfies Harris condition. See Asmussen (1987). The dam model can be generalized to other models for storages and inventories.

There are two problems, related to this thesis, on which we will do future research. The first one is to complete the details in the proof of asymptotic normality of the estimators when $n_0 > 1$ in Definition 1.3.5. As it was said in section 2.3, we need to check a Lindeberg condition for a double array of 1-dependent random variables.

The second problem is concerning to the choosing of the bandwidth. There are several results for the i.i.d. case, some of them are cited in section 2.4. We will start studying this problem trying to extend those results to our case. We also have in mind an iterative method.

All the results above are gotten or will be gotten for real-valued Harris chains. The next step could be extend them for Harris chain with state space $S = R^m$ for some $m \geq 2$. We have not done anything in this direction but we believe that by using the regeneration techniques we can prove similar results.

BIBLIOGRAPHY

S. Asmussen. *Applied Probability and Queues.* John Wiley and Sons Ltd., New York, 1987.

K. B. Athreya and P. Ney. *A new approach to the limit theory of recurrent Markov chains.* Trans. Amer. Math. Soc. 245 (1978) 493-501.

K. B. Athreya, D. McDonald, and P. Ney. *Limit theorems for semi-Markov processes and renewal theory for Markov chains.* Ann. Probability 6 (1978) 788-797.

K. B. Athreya and S. G. Pantula. *Mixing properties of Harris chains and autoregressive processes.* J. Appl. Prob. 23 (1986) 880-892

K. B. Athreya and C. D. Fuh. *Bootstrapping Markov chains: countable case.* J. of Stat. Planning and Inference 33 (1992) 311-331.

K. B. Athreya, H. Doss, and J. Sethuraman. *A proof of convergence of the Markov chain simulation method.* To appear in Ann. Statist.

R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability.* John Wiley and Sons Ltd., New York, 1965.

M. S. Bartlett. *The frequency goodness of fit test for probability chains.* Proc. Camb. Phil. Soc. 47 (1951) 86-95.

I. B. Basawa and B. L. S. Prakasa Rao. *Statistical Inference for Stochastic Processes.* Academic Press, New York, 1980.

R. N. Bhattacharya and E. C. Waymire. *Stochastic Processes with Applications.* John Wiley and Sons Ltd., New York, 1990.

A. W. Bowman. *An alternative method of cross-validation for the smoothing of density estimates.* Biometrika 71 (1984) 353-360.

T. Cacoullos. *Estimation of a multivariate density.* Ann. Inst. Statist. Math. 18 (1964) 179-189.

S. T. Chiu.*Bandwidth selection for kernel density estimation.* Ann. Statist. 19 (1991) 1883-1905.

K. L. Chung. *Contributions to the theory of Markov chains.* I. Res. Nat. Bur. Stand. 50 (1953) 203-208.

K. L. Chung. *A Course in Probability Theory.* 2nd ed., Academic Press, Inc., New York, 1974.

C. Derman. *Some asymptotic distribution theory for Markov chain with a denumerable number of states.* Biometrika 43 (1956) 285-294.

L. Devroye. *A Course in Density Estimation.* Progress in Probability and Statistics. Vol 14, Birkhauser, Boston, 1987.

L.Devroye. *On the non-consistency of an estimate of Chiu.* Statistics and Probability Letters 20 (1994) 183-188.

J. L. Doob. *Stochastic Processes.* John Wiley and Sons, Inc, New York, 1953.

S. E. Ethier and T. G. Kurtz. *Markov Processes. Characterization and Convergence.* John Wiley and Sons, Ltd., New York, 1986.

W. Feller. *An Introduction to Probability Theory and its Applications,* Vol. I, John Wiley and Sons, New York, 1966.

W. Feller. *An Introduction to Probability Theory and its Applications,* Vol. II, John Wiley and Sons, New York, 1971.

P. Hall. *Large sample optimality of least squares cross-validation in density estimation.* Ann. Statist. 11 (1983) 1156-1174.

P. Hall and J. S. Marron. *Extent to which least-squares cross-validation minimises integrated square error in nonparametric density estimation.* Probab. Theory Related Fields 74 (1987) 567-581.

T. E. Harris. *The existence of stationary measures for certain Markov processes.* Proc. of the Third Berkeley Symp. on Math. Stat. and Prob. Vol. 2 (1956) 113-124.

Y. Kanazawa. *Hellinger distance and Kullback-Leibler loss for the kernel density estimator.* Statistics and Probability Letters 18 (1993) 315-321.

S. Karlin and H. G. Taylor. *A First Course in Stochastic Processes.* 2nd ed., Academic Press, New York, 1975.

A. F. Karr. *Point Processes and Their Statistical Inference.* 2nd ed., Marcel Dekker, New York, 1991.

S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Springer Verlag, New York, 1993.

P. A. P. Moran. *The Theory of Storage.* Methuen, London, 1959.

E. Nummelin. *A splitting technique for Harris recurrent Markov chains.* Zeitschrift fur Wahr. und verwandte Gebiete 43 (1978) 309-318.

E. Nummelin. *General irreducible Markov chains and non- negative operators.* Camb. Univ. Press, London, 1984.

E. Parzen.*On estimation of a probability density function and mode.* Ann. Math. Statist. 33 (1962) 1065-1076.

M. J. M. Posner and D. Zuckerman.*Semi-Markov shock models with additive damage.* Adv. Appl. Probability 18 (1986) 772-790.

N. U. Prabhu. *Stochastic Storage Processes. Queues, Insurance, and Dams.* Springer Verlag, N.York. (1980).

B. L. S Prakasa Rao. *Nonparametric Functional Estimation.* Academic Press, New York, 1983.

D. Revuz. *Markov Chains.* North-Holland, Amsterdam, 1975.

M. Rosenblatt. *Remarks on some nonparametric estimates of a density.* Ann. Math. Statist. 27 (1956) 832-835.

M. Rosenblatt. *Density estimates and Markov sequences.* In Nonparametric Techniques in Statistical Inference, Ed by M. Puri, (1970) 199-210.

M. Rosenblatt. *Markov Processes. Structure and Asymptotic Behavior.* Springer Verlag, New York, 1971.

G .G . Roussas. *Nonparametric estimation in Markov processes.* Annals of Inst. of Stat. Math. (1969a) 73-87.

G. G. Roussas. *Nonparametric estimation of the transition distribution function of a Markov process.* Annals of Math. Stat. (1969b) 1380-1400.

G. G. Roussas and D. Ioannides. *Moment inequalities for mixing sequences of random variables.* Stochastic Annal. Appl. 5(1), (1987) 61-120.

G. G. Roussas. *Estimation of transition distribution function and its quantiles in Markov processes:Strong consistency and asymptotic normality.* In Nonparametric Functional Estimation and Related Topics. Ed by G. G. Roussas (1991a) 443-462.

G. G. Roussas. *Recursive estimation of the transition distribution function of a Markov process:* Asymptotic normality. Statistics and Probability Letters 11 (1991b) 435-447.

M. Rudemo. *Empirical choice of histograms and kernel density estimators.* Scand. J. Statist. 9 65-78.

W. Rudin. *Real and Complex Analysis.* Third edition, McGraw-Hill, New York, 1987.

D. Scott and G. R. Terrel. *Biased and unbiased cross-validation in density estimation.* J. Amer. Statist. Assoc. 82 (1987) 1131-1146.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London, 1986.

W. Stadje. *A New Look at The Moran Dam.* J. Appl. Prob. 30 (1993) 489-495.

C. J. Stone. *An asymptotically optimal window selection rule for kernel density estimates.* Ann. Statist. 12 (1984) 1285-1297.